

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LE PARTAGE DE DONNÉES SCIENTIFIQUES : ANALYSE DES RÉGIMES D'ÉCHANGE  
DANS UNE COMMUNAUTÉ DE RECHERCHE SUR L'ARCTIQUE

MÉMOIRE  
PRÉSENTÉ  
COMME EXIGENCE PARTIELLE  
DE LA MAÎTRISE EN COMMUNICATION

PAR  
OLIVIER GRATTON-GAGNÉ

JUIN 2011

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

## REMERCIEMENTS

Les deux dernières années sont parmi les plus belles de ma vie. La maîtrise en communication et le processus de rédaction du mémoire sont des étapes exigeantes, mais non sans côtés positifs. Je tiens donc à remercier tous ceux qui m'ont appuyé d'une manière ou d'une autre dans les dernières années.

Du côté universitaire, je dois d'abord toute ma gratitude à ma directrice Florence Millerand, qui a su me motiver, m'encadrer et me supporter dans mes projets, et ce avant même que je finisse mon baccalauréat. Je me considère très chanceux d'avoir travaillé avec une personne aussi professionnelle et intègre que Florence.

Également, ce projet ne serait pas le même sans le soutien d'*ArcticNet* pour notre entreprise. Je tiens à remercier cette organisation, et plus particulièrement Christine Barnard, Josée Michaud et Warwick Vincent pour leur accueil chaleureux.

Mon passage à la maîtrise n'aurait pas été le même sans le Laboratoire de communication médiatisé par ordinateur et le Centre interuniversitaire de recherche sur la science et la technologie. Je remercie l'équipe d'employés du Centre interuniversitaire de recherche sur la science et la technologie (CIRST) pour l'excellent travail qu'ils font tous les jours. Au Laboratoire de communication médiatisée par ordinateur (LabCMO), un grand merci à tous mes collègues : vous avez su m'inspirer et me motiver à aller jusqu'au bout de mon projet.

Mes collègues de maîtrise ont été une source de support et d'inspiration tout au long de ma recherche. Je me dois de vous remercier chaleureusement pour tous ces beaux moments passés à l'Université et à l'extérieur de l'Université.

Je me dois aussi de souligner l'apport de Florence Millerand, de Serge Proulx et de Pierre-Léonard Harvey, les membres du jury de ce mémoire. Leurs commentaires judicieux et leurs encouragements lors de la présentation du projet de mémoire m'ont été très utiles à toutes les

étapes de la rédaction. Je ne peux pas passer sous silence le dévouement de Pierre Mongeau et Isabelle Mahy, qui ont régulièrement fait parvenir des lettres de recommandations pour soutenir ma candidature dans des concours de bourse.

Un grand merci à Rosalie Dion-Picard, Hélène Gratton, Sarah Landry, Marie Léger-St-Jean et Andrée-Anne Marsan qui ont relu mon mémoire attentivement et m'ont souligné les passages à travailler et les coquilles à corriger.

Du côté personnel, je tiens à remercier ma copine Andrée-Anne, qui a été avec moi durant toutes les étapes de la maîtrise. Son soutien est inestimable. Également, je me dois de remercier mes parents François Gagné et Hélène Gratton, qui m'ont appuyé sans fautes avant (23 ans!) et pendant ma maîtrise. Je vous remercie tous du fond du cœur.

- - - - -

## TABLE DES MATIÈRES

LISTE DES FIGURES .....	ix
RÉSUMÉ .....	xi
INTRODUCTION .....	1
CHAPITRE I	
PROBLÉMATIQUE.....	3
1.1 L'omniprésence des technologies de l'information.....	3
1.2 La collaboration scientifique en ligne : typologies et définitions.....	4
1.3 L'échange de données scientifiques par les bases de données en ligne.....	5
1.4 Les bases de données dans la sphère publique .....	6
1.4.1 Les médias et les bases de données .....	6
1.4.2 Le Web social et le partage de données.....	7
1.5 Les promesses des bases de données en science.....	8
1.6 Les enjeux des bases de données .....	10
1.7 Les bases de données comme objet d'étude .....	12
1.7.1 Le portrait global des bases de données en science.....	12
1.7.2 La propriété intellectuelle.....	13
1.7.3 Les fonctions des données en science.....	13
1.7.4 Les usages des bases de données .....	14
1.8 Les questions de recherche .....	15
1.9 La pertinence communicationnelle.....	16
CHAPITRE II	
CADRE THÉORIQUE.....	19
2.1 Les questions métathéoriques .....	19
2.1.1 L'ontologie .....	19
2.1.2 L'épistémologie.....	20

2.1.3	L'axiologie.....	21
2.2	Les grandes perspectives théoriques sur la science et la technologie .....	22
2.2.1	La « nouvelle » sociologie des sciences et des technologies .....	22
2.2.2	Les études en informatique sociale .....	23
2.2.3	La sociologie des usages.....	24
2.3	Une piste pour saisir les modalités d'usage : les régimes d'échange.....	28
2.3.1	Les régimes d'action .....	28
2.3.2	Les régimes pour comprendre les communautés du logiciel libre et les <i>wikis</i> ...	29
2.3.3	Les régimes de communication scientifique.....	29
2.3.4	Les régimes d'échange.....	30
2.3.5	La pertinence de la théorie des régimes d'échange .....	31
2.3.6	Les définitions des concepts utilisés.....	33
CHAPITRE III		
	CADRE MÉTHODOLOGIQUE.....	35
3.1	Le terrain de recherche.....	35
3.1.1	Une présentation d' <i>ArcticNet</i> .....	35
3.1.2	La pertinence du terrain .....	37
3.2	L'orientation méthodologique .....	38
3.3	Les considérations éthiques .....	38
3.4	La cueillette des données .....	40
3.4.1	Les outils de cueillette .....	40
3.4.2	Les entrevues semi-dirigées.....	40
3.4.3	Le déroulement de la collecte de données .....	42
3.5	L'analyse des données .....	44
3.5.1	L'approche générale .....	44
3.5.2	Le déroulement de l'analyse.....	46
CHAPITRE IV		
	RÉSULTATS .....	51
4.1	Le contexte général du partage de données .....	51
4.1.1	Le contexte de la recherche sur l'Arctique .....	52
4.1.2	Les différences dans la nature même des données.....	53
4.1.3	La collaboration multidisciplinaire.....	54

4.1.4	Une perception commune : le développement du partage de données est inéluctable.....	55
4.2	Le régime de coopération .....	56
4.2.1	La proximité des acteurs.....	57
4.2.2	L'importance de la confiance .....	58
4.2.3	La facilité des échanges.....	58
4.2.4	Une réciprocité appréciée, mais facultative.....	59
4.3	Le régime de don .....	60
4.3.1	Des partenaires distants .....	61
4.3.2	La réciprocité des dons .....	61
4.3.3	Les étudiants comme intermédiaires .....	63
4.4	Le régime de sous-traitance .....	63
4.4.1	Des partenaires hétérogènes .....	64
4.4.2	La spécialisation des tâches entre chercheurs.....	65
4.4.3	L'encadrement formel des échanges .....	66
4.5	Au-delà des régimes .....	67
4.5.1	Le refus de partager .....	67
4.5.2	Les sanctions aux mauvais collaborateurs.....	68
CHAPITRE V		
DISCUSSION .....		69
5.1	L'importance des relations de collaboration.....	69
5.1.1	La condition au partage : une relation de collaboration .....	70
5.1.2	Les technologies et les relations de collaboration .....	70
5.1.3	L'apport de l'informatique sociale .....	71
5.1.4	Deux exemples d'intégration de la relation au dispositif .....	72
5.2	Une nouvelle figure : le chercheur-contributeur.....	74
5.2.1	Comprendre le contexte et la qualité des données.....	74
5.2.2	Le chercheur-contributeur .....	75
5.2.3	Le chercheur-contributeur, figure du Web social? .....	76
5.2.4	Un contributeur invisible? .....	77
5.3	Repenser les régimes d'échange.....	78
5.3.1	Une vision trop optimiste du partage de données .....	78

5.3.2	Pour un modèle dynamique des régimes d'échange .....	80
5.3.3	Le passage d'un régime d'échange à un autre .....	80
CONCLUSION .....		85
6.1	Le partage de données et la collaboration scientifique .....	85
6.1.1	La pluralité des régimes d'échange .....	85
6.1.2	Avant le partage, il y a une relation de collaboration .....	86
6.1.3	Le chercheur-contributeur, une nouvelle figure à explorer.....	86
6.2	Les limites de la recherche.....	87
6.2.1	Le biais positif de la théorie des régimes d'échange .....	88
6.2.2	Les limites de nos appuis empiriques .....	88
6.3	Les pistes de recherches futures.....	89
6.3.1	Les nouvelles formes de travail scientifique.....	89
6.3.2	L'étude quantitative du partage de données.....	90
6.3.3	Développer la théorie des régimes d'échange .....	91
6.3.4	Agir et réfléchir sur les dispositifs de collaboration .....	91
BIBLIOGRAPHIE .....		93

- - - - -



## LISTE DES FIGURES

Figure 3.1	Nombre total de codes et nombre d'entrevues codées .....	47
Figure 5.1	Gradation des régimes selon le niveau de confiance.....	82



## RÉSUMÉ

Ce mémoire porte sur le partage de données dans la recherche sur l'Arctique. Le partage de données est ici problématisé dans le contexte du développement d'infrastructures scientifiques de collaboration à grande échelle. Notre recherche est structurée autour de la question centrale suivante : comment rendre compte de la pluralité et de la complexité des usages des bases de données scientifiques? Pour ce faire, nous mobilisons les traditions théoriques des études en science, technologie et société, en informatique sociale et en sociologie des usages. Plus précisément, nous recourons à la théorie des régimes d'échange. Cette théorie propose de considérer les pratiques de partage de données à travers trois régimes d'échanges distincts, soit le régime de coopération, le régime de don et le régime de sous-traitance. Notre méthodologie est basée sur une approche qualitative et ascendante et notre terrain de recherche est le réseau de recherche *ArcticNet*. Nos résultats illustrent la variété et la subtilité des modalités d'échange de données et mettent au jour des enjeux du travail scientifique actuel. Nous terminons ce mémoire en proposant le concept du chercheur-contributeur. Selon cette hypothèse, les scientifiques occupent de plus en plus les fonctions des professionnels des sciences de l'information à mesure que les échanges de données s'accélèrent.

**Mots clés :** données, bases de données, recherche scientifique, Arctique, partage, infrastructures scientifiques, régime d'échange

--

--

--

--

## INTRODUCTION

Depuis plusieurs décennies, l'activité scientifique est fortement marquée par la collaboration. Les chercheurs travaillent souvent au sein d'une équipe, publient des communications à plusieurs auteurs et participent à des colloques qui réunissent des milliers de personnes. Avec le développement rapide des technologies de l'information et de communication, beaucoup soutiennent qu'il est maintenant possible de pousser la coopération scientifique encore plus loin.

Le présent mémoire se penche sur une des facettes de la collaboration scientifique, à savoir le partage de données par des bases de données en ligne dans le réseau *ArcticNet*.

Dans le chapitre I, nous présentons le contexte général du développement d'infrastructures scientifiques et nous cernons notre objet de recherche. Subséquemment, nous dressons un portrait des recherches antérieures sur ce sujet pour ensuite énoncer nos questions de recherche. Le chapitre II se consacre à l'explicitation des bases théoriques sur lesquelles notre projet se fonde. Nous y formulons d'abord nos a priori épistémologiques pour ensuite décrire les trois champs d'étude que nous mobilisons pour cette recherche. Le chapitre II se termine par la présentation de notre principal outil d'analyse : la théorie des régimes d'échange (Milanovic, Pontille et Cambon-Thomsen, 2007). Le chapitre III a pour objectif de présenter notre démarche méthodologique. Nous y décrivons d'abord notre terrain de recherche : *ArcticNet*, un réseau regroupant une centaine de chercheurs spécialisés dans l'Arctique canadien. Le reste du chapitre est consacré à la description de la démarche de cueillette de données et d'analyse de celle-ci.

Les résultats de notre recherche sont présentés dans le chapitre IV. Nous y décrivons d'abord les principaux éléments contextuels de la recherche en Arctique. Ensuite, les pratiques d'échange de données y sont détaillées, révélant la grande variété des arrangements normatifs et institutionnels qui supportent ces collaborations. La fin du chapitre IV est consacrée aux résultats qui ne pouvaient pas être directement liés au modèle des régimes d'échange. Dans le chapitre V,

nous discutons ces résultats à la lumière des ancrages théoriques sollicités dans ce mémoire pour développer une série d'hypothèses sur la collaboration scientifique. Enfin, dans la conclusion, nous évoquons les idées majeures défendues par ce mémoire et mettons en lumière les pistes de recherches futures qu'il importe de considérer.

---

## CHAPITRE I

### PROBLÉMATIQUE

#### 1.1 L'omniprésence des technologies de l'information

Les dernières décennies ont vu se développer le domaine des technologies de l'information et de la communication (TIC). Objets de fascination, de critiques et de controverses, les TIC ont gagné en importance dans une grande variété de milieux. L'informatisation des organisations a changé de manière importante la nature et les conditions du travail et des pans entiers de l'économie sont désormais consacrés à la production de biens et de services liés aux TIC. Ces développements amènent aussi de nouveaux enjeux, tels que le remplacement à grande échelle de travailleurs humains par des dispositifs techniques ou encore « cette volonté de transférer toujours plus avant les fonctions intellectuelles de l'homme vers la machine [...] » (Breton, 1987: 143).

Loin de se limiter à la sphère professionnelle, les TIC ont également pris leur place dans les foyers, avec la mise en marché des ordinateurs personnels, des téléphones portables, etc. Plus que de simples outils de travail, les dispositifs techniques se déploient maintenant sur presque toutes les facettes de l'activité humaine. Certains sont utilisés dans la sphère intime ou dans le divertissement, alors que d'autres servent à des fins d'expression politique ou artistique. Plus récemment, on remarque que les frontières entre les espaces privés, personnels et professionnels se brouillent, notamment par la très grande disponibilité des technologies et des connexions au réseau Internet (sans fil, par téléphone portable « intelligent », etc.). Bref, la question technologique est rendue ubiquiste dans les sociétés occidentales.

Comme dans les sphères professionnelles et domestiques, les avancées des TIC et le foisonnement des usages qui les accompagnent ont eu un effet certain dans le milieu de la recherche scientifique. Entre autres, les développements informatiques des dernières décennies

ont permis de mettre en place des outils de mesure, de stockage et d'analyse des données toujours plus sophistiqués. Si l'on reconnaît aisément la place centrale des dispositifs informatiques dans les disciplines usant intensément des mathématiques (physique, statistiques, économie, etc.), des disciplines telles que l'anthropologie, la philosophie ou les études littéraires voient leurs pratiques changer avec l'introduction des ordinateurs (Katzen, 1986).

Avec la mise en place des réseaux informatiques tels que l'Internet, les dispositifs informatiques sont amenés à jouer de nouveaux rôles. Aux fonctions de calcul et de traitement de l'information s'ajoute un éventail d'outils de communication et de collaboration qui viennent brouiller les frontières physiques et temporelles des laboratoires et des universités. Qu'on les appelle « collaboratoires », « cyberinfrastructures » ou « e-Science », ces dispositifs ont en commun de permettre à un nombre relativement élevé de chercheurs de partager des informations et d'accéder à des ressources communes comme des bases de données partagées, des architectures de calcul distribué ou « grid », des outils de visualisation et de modélisation, etc. Dans diverses communautés scientifiques, beaucoup soutiennent que ces infrastructures de collaboration vont permettre à l'activité scientifique de « changer d'échelle » et de produire de nouvelles découvertes (Arzberger *et al.*, 2004b, 2004a; Corti, 2000; Gilbert, 1991; Koslow, 2000; National Research Council, 1997; Schroder, 2003; The Governing Council of the Organization for Human Brain Mapping, 2001). Dans cette perspective, des projets de « cyberinfrastructures » sont en cours pour soutenir l'activité scientifique à grand déploiement (Atkins, 2003; Wouters et Reddy, 2003).

## 1.2 La collaboration scientifique en ligne : typologies et définitions

D'emblée, il importe de reconnaître que les infrastructures de collaboration scientifique sont très variées tant dans leurs formes, leurs fonctions que leurs échelles. En effet, comment peut-on comparer un accélérateur de particules haut de plusieurs étages à un simple ordinateur partagé dans un laboratoire? Pour guider notre réflexion, voyons les distinctions et les typologies qui ont été proposées.

D'abord, Wouters *et al.* (2007: 325) font la différence entre cyberinfrastructures et e-Science. Le premier terme, plus commun aux États-Unis, renvoie aux conditions matérielles et techniques de la science. Le second terme, plus utilisé en Europe, caractérise les pratiques scientifiques à l'ère du numérique. Pour Wouters *et al.*, il demeure que ces deux vocables réfèrent



à l'idée d'une nouvelle forme de recherche scientifique caractérisée par son usage intensif d'outils informatiques.

Ensuite, Bos *et al.* (2007) proposent une définition du collaboratoire. Si les termes de cyberinfrastructures et d'e-Science restent assez généraux, le collaboratoire renvoie déjà à un objet plus précis :

*[il s'agirait d'une] entité organisationnelle qui se déploie sur de grandes distances, qui supporte des interactions humaines riches et récurrentes orientées vers un domaine de recherche commun, qui développe les liens entre les chercheurs qui se connaissent et qui ne se connaissent pas et qui offre l'accès à des données, des artefacts et des outils nécessaires à l'accomplissement des activités de recherche. (Bos et al., 2007)*

Bos *et al.* établissent une taxonomie de sept types de collaboratoires. Les types relevés distinguent les collaborations autour des instruments (comme les accélérateurs de particules ou des télescopes communs), des données (comme les bases de données) et des savoirs (comme les salles de classe en ligne ou des laboratoires virtuels). Dans le cadre de ce mémoire, nous portons notre attention sur le cas des bases de données.

### 1.3 L'échange de données scientifiques par les bases de données en ligne

Simplement définie, une base de données est un dispositif informatique matériel ou logiciel qui permet le dépôt, l'archivage, l'organisation, l'accès et le traitement d'informations et de données<sup>1</sup>. En science, les bases de données sont très fréquemment utilisées à l'échelle locale pour la manipulation et l'analyse des données de recherche. Avec l'accès de plus en plus facile aux réseaux informatiques, des bases de données *en ligne* sont développées. Ces dernières, hébergées sur des serveurs distants, permettent à un grand nombre de collaborateurs géographiquement dispersés d'échanger leurs données par l'intermédiaire du réseau.

---

<sup>1</sup> Par données, nous entendons : « n'importe quelle information qui peut être stockée sous une forme numérique, incluant textes, nombres, images, vidéos et films, audio, logiciels, algorithmes, équations, animations, modèles, simulations, etc. » (National Science Board, 2005).

Une distinction importante doit être faite entre les bases de données de documents finis (par exemple, des articles de revues scientifiques ou des encyclopédies) et les bases de données qui rassemblent des données primaires de recherche (par exemple, des taux de cholestérol en sciences de la santé ou des mesures de diamètres d'arbres en écologie). Alors que les premières sont des extensions numérisées de documents imprimés et constituent l'aboutissement du cycle de recherche, les secondes rassemblent des données qui sont habituellement conservées par les chercheurs et qui sont au cœur du processus de recherche. Dans la présente recherche, c'est sur les bases de données primaires que nous nous penchons.

#### 1.4 Les bases de données dans la sphère publique

##### 1.4.1 Les médias et les bases de données

Loin de rester confinés aux universités et aux laboratoires, les promesses et les espoirs portés par les projets de grandes bases de données trouvent des échos dans la sphère publique. Très récemment, *The Economist* titrait son édition du 27 février 2010 « *The data deluge. And how to handle it* » en plus de consacrer une section de quatorze pages sur la problématique des données scientifiques, commerciales et gouvernementales toujours plus abondantes (Anonyme, 2010).

Pour *The Economist*, la prolifération des données et le coût dérisoire de stockage de celles-ci va provoquer des changements importants dans le travail scientifique. Dans la même perspective, la revue *Nature* publiait un éditorial sur le fait que le travail des scientifiques ressemble de plus en plus à du développement de logiciels et d'algorithmes (Gilbert, 1991). Une fois conçus, ces programmes informatiques n'auraient qu'à être « lancés » dans une mer de données pour en ressortir avec des résultats. De telles pratiques sont déjà à l'œuvre dans des disciplines telles que l'astronomie ou la physique nucléaire, dans lesquelles la quantité de données collectées dépassent l'entendement (Anonyme, 2010). Selon *The Economist*, ce changement radical des méthodologies scientifiques s'observe également dans le monde des affaires et dans les différents appareils gouvernementaux. La revue britannique rappelle que ce déluge de données comporte évidemment sa part de risques, avec en premier lieu la sécurité des données

confidentielles<sup>2</sup> et la capacité limitée des infrastructures publiques et privées pour absorber la quantité sans cesse grandissante d'information.

#### 1.4.2 Le Web social et le partage de données

Les nombreuses discussions que suscitent les bases de données dans l'actualité se manifestent aussi dans la problématique émergente du *Web social* (Millerand, Proulx et Rueff, 2010).

Pour Millerand, Proulx et Rueff (2010), le vocable Web social désigne deux phénomènes distincts, mais liés. D'une part, le Web social évoque le développement des dispositifs, des logiciels et des infrastructures d'Internet survenu après l'éclatement de la bulle boursière des entreprises *point com*. Regroupée sous l'étiquette *Web 2.0*, une pléthore de sites et d'applications s'est mise en place pour fournir aux usagers des plateformes pour créer, partager, commenter ou indexer une grande variété de contenus (O'Reilly, 2005). Du partage de vidéos aux encyclopédies contributives, les multiples avatars du Web 2.0 ont pour point commun de placer l'utilisateur et ses contributions au cœur de leur fonctionnement. D'autre part, le Web social désigne les nouveaux usages qui ont émergé en concomitance avec les innovations techniques du Web 2.0. S'affranchissant graduellement d'un rôle passif de consommateur, l'utilisateur du Web social est appelé à générer des contenus et à les partager. Alimenté par les mouvements du logiciel libre, la distinction entre concepteur et usager se brouille et l'on invite les utilisateurs à contribuer au développement des dispositifs eux-mêmes.

Au-delà de la vogue actuelle de ces technologies, certaines innovations du Web social interrogent directement les manières traditionnelles de générer, d'organiser et de diffuser la connaissance (Auray *et al.*, 2009). Dans son essai *Everything is miscellaneous*, Weinberger (2007) soutient que les systèmes traditionnels d'organisation et de conservation des savoirs (les

---

<sup>2</sup> À cet effet, le Royaume-Uni fut secoué par une série de scandales entourant la perte de données par différentes branches de son gouvernement. De novembre 2007 à janvier 2008, le service d'impôt, le ministère de la Santé et le ministère de la Défense ont successivement perdu des quantités importantes de données confidentielles à la suite de négligences de leurs employés (Anonyme, 2008; Chivers, 2007; Summers, 2007).

bibliothèques, les archives, les livres, les médias imprimés, etc.) sont sur le point d'être renversés par l'enchevêtrement anarchique des références, des liens, des commentaires, des taxonomies et des préférences que les internautes créent chaque jour. De ce fouillis informationnel émergerait un nouvel ordonnancement en perpétuelle évolution.

Des activistes du Web social ont multiplié les appels pour diffuser les données détenues par les gouvernements et les organisations internationales afin qu'un grand nombre d'utilisateurs puissent les traiter, les organiser, les diffuser, etc. Parmi ceux-ci, le spécialiste en santé internationale Hans Rosling a mis au point le site Internet *Gapminder*, lequel permet aux utilisateurs de générer des graphiques et des visualisations à loisir à partir de données de l'ONU et d'autres organisations (Radhakrishna, 2010).

Sans surprise, plusieurs projets récents d'infrastructures de collaboration scientifique reprennent les idéaux de partage et de contribution à grande échelle telle qu'ils se manifestent dans le Web social. À cet effet, Burk (2007) remarque que l'éthos scientifique tel que défini par Merton (1974) ressemble grandement aux normes présentes dans les communautés du logiciel libre. Conséquemment, plusieurs projets d'infrastructures de partage prévoient utiliser les licences et les dispositifs techniques et légaux issus du monde du logiciel libre et du Web social (licence publique générale GNU, *Creative Commons*, etc.) pour assurer un partage gratuit et sans encombre.

### 1.5 Les promesses des bases de données en science

Le pari des promoteurs des bases de données est le suivant : en partageant à grande échelle les données qui restaient auparavant confinées dans les murs des laboratoires, nous serons en mesure d'accéder à une quantité inégalée de données qu'il sera possible d'analyser à loisir. Il s'agirait d'un puissant catalyseur de l'activité scientifique permettant à des chercheurs de plusieurs pays et de plusieurs disciplines de mettre en commun leurs efforts et de relever les grands défis scientifiques de notre époque (climat, environnement, etc.) (Arzberger *et al.*, 2004a; Bowker, 2006).

Au courant des années 1990 et 2000, plusieurs acteurs du monde scientifique ont lancé des appels pour la mise en place de grandes bases de données et pour inciter les chercheurs à rendre leurs données publiques à l'aide de ces dispositifs (Anonyme, 2000a, 2000b; Arzberger *et al.*, 2004b, 2004a; Casey, 2003; Corti, 2000; Drake, 2003; Gilbert, 1991; Koslow, 2000; National

Research Council, 1997; Schroder, 2003; The Governing Council of the Organization for Human Brain Mapping, 2001; Wellcome Trust, 2003). Ces appels mettent de l'avant plusieurs des promesses et des potentialités des bases de données. Sans être exhaustive, la liste suivante résume les principaux bienfaits anticipés des bases de données en ligne.

- Une ressource centralisée au sein d'une communauté scientifique permettrait d'éviter les doublons (Corti, 2000). Avant d'aller collecter lui-même ses données sur le terrain, un chercheur pourrait bénéficier de données déjà existantes et investir son temps, ses fonds et ses efforts de recherche à de meilleures fins.
- Le développement de bases de données scientifiques amènerait la communauté scientifique à se doter de standards communs (Beaulieu, 2001, 2003). Les interfaces de partage qu'offrent les bases de données conduiraient la communauté à adopter un langage plus uniforme, des unités standardisées et des procédures expérimentales normalisées.
- Le fait d'exposer leurs données primaires contraindrait les scientifiques à une plus grande rigueur à toutes les étapes de leur démarche (Anonyme, 2000a; Corti, 2000).
- Grâce aux bases de données, les chercheurs pourraient produire de nouvelles recherches à partir d'anciennes données (Schroder, 2003; Zimmerman, 2007, 2008). Par exemple, des données climatiques pourraient servir des recherches en écologie et vice-versa.
- Les bases de données permettraient à la recherche scientifique de changer d'échelle, d'atteindre une plus grande envergure. Pour les disciplines telles que la génomique, la protéomique, la neuroinformatique ou l'étude de la biodiversité, le recours à de grandes quantités de données serait désormais inséparable de la pratique scientifique quotidienne (Bowker, 2000; Brown, 2003; Keil, 1990). Aussi, des études comparatives et des méta-analyses pourraient bénéficier d'un plus grand partage des données (Anonyme, 2000a; Corti, 2000).
- Les investissements publics seraient plus efficaces si les données de recherche deviennent accessibles à toute la communauté scientifique (Arzberger *et al.*, 2004a; Schroder, 2003; Wellcome Trust, 2003). Dans leurs rôles de promoteur de la recherche scientifique et du savoir collectif, les agences de financement pourraient s'assurer que

les recherches subventionnées aient un maximum de retombées dans leur milieu en établissant des infrastructures de partage des données.

- L'établissement de bases de données centrales rendrait accessibles les données, les infrastructures et les ressources scientifiques des pays développés aux chercheurs des pays en développement (Luo et Olson, 2008; National Research Council, 1997).
- Le déploiement d'outils de collaboration serait envisagé comme l'occasion de passer d'une petite science (*little science*) marquée par des pratiques locales et dispersées, à une grande science (*big science*) où un grand nombre de chercheurs collaboreraient autour d'instrumentations complexes (Borgman, Wallis et Enyedy, 2007; Wouters et Reddy, 2003).

#### 1.6 Les enjeux des bases de données

Comme d'autres dispositifs techniques, les bases de données en ligne suscitent beaucoup d'espoir. Pourtant, il a été démontré que les dispositifs seuls ne garantissent pas leur appropriation effective par les usagers (Jouët, 2000; Kling, 1999). Comme Olson et Olson (2000) l'ont démontré, les technologies de collaboration médiatisée parviennent difficilement à franchir les distances physiques, temporelles ou culturelles qui séparent les usagers. De plus, on remarque finalement que peu de ces outils de collaboration ont remporté de véritables succès dans le milieu de la recherche (Bos *et al.*, 2007). Au terme d'une recension des écrits sur le sujet, nous avons par ailleurs établi que de nombreux enjeux nuisaient à l'adoption de pratiques de partage de données au sein des communautés scientifiques, comme la compétition entre les chercheurs, le manque de standards, les questions éthiques et légales, etc. (Gratton-Gagné, 2009; Gratton-Gagné et Millerand, à paraître).

Parmi les enjeux importants liés au développement et aux usages des bases de données, notons que les études comparatives laissent entrevoir un fossé grandissant entre les disciplines qui ont massivement recours à ces outils et celles qui tardent à les adopter (Carlson et Anderson, 2007; Talja et Maula, 2003). Évidemment, les bases de données ne se prêtent pas à toutes les pratiques scientifiques. Cependant, dans un contexte où les gouvernements et les agences de financement sont interpellés à promouvoir et à soutenir le partage des données scientifiques (Arzberger *et al.*, 2004b, 2004a; Drake, 2003; Wellcome Trust, 2003), il est à craindre que plusieurs communautés scientifiques ne puissent pas bénéficier des mêmes subventions et des

mêmes infrastructures dont jouissent des disciplines plus *branchées*. *A fortiori*, les disciplines qui auront incorporé le partage de données dans leurs pratiques seront mieux en mesure de plaider pour l'utilité publique de leurs projets scientifiques et ainsi récolter d'autres subventions.

S'il existe un fossé dans le recours aux bases de données, c'est entre les méthodes qualitatives et les méthodes quantitatives (Carlson et Anderson, 2007; Corti, 2000). En effet, par nature, les données quantitatives sont plus facilement numérisées et échangées d'un chercheur à un autre. *A contrario*, les données qualitatives sont riches, complexes et hautement contextuelles. Par exemple, Carlson et Anderson (2007) relatent comment des anthropologues peinent à cataloguer efficacement leur matériel de recherche parce que celui-ci est trop disparate et contextuel. De plus, les deux auteurs soulignent les nombreux enjeux éthiques entourant la numérisation de documents portant sur des sujets humains. Par exemple, est-il possible et souhaitable d'obtenir le consentement des sujets impliqués pour un partage libre des données de recherche?

Selon Berg et Goorman (1999), plus une information est appelée à circuler, plus des efforts importants doivent être accomplis pour la dénouer de son contexte, notamment par l'inscription de métadonnées<sup>3</sup>. Dans le cas des bases de données, les scientifiques qualitatifs doivent donc investir des efforts considérables pour affranchir leurs données et « capturer » les éléments contextuels pertinents. Alors que ces efforts sont inhérents à toute démarche quantitative, ils prennent la forme d'une tâche supplémentaire et se placent souvent en opposition directe avec les ancrages méthodologiques et épistémologiques des sciences qualitatives.

Une autre problématique entourant l'usage de bases de données est celle des relations Nord-Sud. Bien que Luo et Olson (2008) soulignent les réussites de laboratoires réunissant des scientifiques des pays développés et des pays en développement, on remarque que les infrastructures souvent défaillantes des pays du Sud rendent le partage au sein de ces dispositifs plus difficile. De même, les fonds de recherche limités et les problèmes de visas empêchent souvent les chercheurs des pays en développement de participer aux rencontres en face à face

---

<sup>3</sup> Par métadonnées, nous entendons « la documentation des données [...] un niveau d'information supérieur ou des instructions qui décrivent le contenu, le contexte, la qualité, la structure et l'accessibilité d'un ensemble spécifique de données » (Michener *et al.*, 1997: 331).

entre les membres de ces infrastructures de collaboration. Enfin, les cas étudiés par Luo et Olson montrent que les échanges sont surtout unidirectionnels, les pays du nord fournissant données, expertises et ressources au pays du sud sans réelle réciprocité, recréant des relations de type colonialistes.

### 1.7 Les bases de données comme objet d'étude

Un nombre grandissant de chercheurs se penchent sur les bases de données comme objet d'étude, à partir de diverses perspectives disciplinaires : les études en science, technologie et société (STS), la bibliothéconomie, la sociologie, le droit, la communication, les études littéraires, etc.

Certains auteurs ont cherché à identifier les enjeux principaux des bases de données scientifiques (Arzberger *et al.*, 2004b, 2004a; Casey, 2003; Hey, Tansley et Tolle, 2009; National Research Council, 1997; Olson, Zimmerman et Bos, 2008; Schroder, 2003; Wouters et Reddy, 2003). D'autres se sont penchés sur les questions juridiques du partage des données (Burk, 2007; Eckersley *et al.*, 2003). Des chercheurs plus près des traditions interprétatives ont tenté de mettre en évidence l'enracinement social des données de recherche et de leurs infrastructures (Birnholtz et Bietz, 2003; Bowker, 2000; Carlson et Anderson, 2007; Coopmans, 2006). Enfin, plusieurs auteurs ont cherché à rendre compte des usages des bases de données dans le contexte scientifique (Beaulieu, 2003; Borgman, Wallis et Enyedy, 2007; Brown, 2003; Hine, 2006; Pontille *et al.*, 2006; Star et Ruhleder, 1994, 1996; Zimmerman, 2007, 2008). De cette dernière avenue de recherche, la notion de régime a émergé pour comprendre la pluralité et la complexité des pratiques observées autour des bases de données (Bohlin, 2004; Hilgartner, 1995; Milanovic, Pontille et Cambon-Thomsen, 2007). Dans le cadre de notre mémoire, le concept de régime est mobilisé pour développer notre compréhension des usages des bases de données. Cependant, avant de développer plus en détail notre cadre théorique, nous exposerons brièvement les différentes perspectives de recherche adoptées pour l'étude des bases de données en science.

#### 1.7.1 Le portrait global des bases de données en science

Une première piste de recherche adoptée pour l'étude des bases de données scientifiques procède à une analyse globale et cherche à dégager les grands enjeux et les principaux obstacles à l'appropriation de ces dispositifs techniques. De ces textes, on peut voir émerger l'idée centrale selon laquelle « les données de recherches financées par le public doivent être les plus accessibles



possible » (Arzberger *et al.*, 2004b: 136). Cependant, il est reconnu que la mise en place d'une telle accessibilité aux données de recherche est sérieusement minée par des obstacles technologiques, organisationnels, budgétaires, légaux et culturels (Arzberger *et al.*, 2004b, 2004a).

Dans le même ordre d'idée, deux livres majeurs sur la question des bases de données et de la collaboration scientifique en ligne ont été publiés récemment. *Scientific Collaboration on the Internet* (Olson, Zimmerman et Bos, 2008) et *The Fourth Paradigm : Data-Intensive Scientific Discovery* (Hey, Tansley et Tolle, 2009), deux ouvrages collectifs, proposent plusieurs études de cas ainsi que des pistes d'analyse et de théorisation pour comprendre les pratiques de collaboration scientifique par les TIC. Adoptant un regard près des études en STS, les auteurs de *Scientific Collaboration* soulignent les potentialités des infrastructures techniques, mais restent lucides quant à leurs limites et leurs enjeux. De son côté, *The Fourth Paradigm* aborde ces questions avec une optique plus proche de l'ingénierie informatique (Gray *et al.*, 2009).

#### 1.7.2 La propriété intellectuelle

Ces travaux s'axent autour des questions juridiques liées à la propriété intellectuelle. De ces écrits, il ressort que les bases de données scientifiques sont prises dans une double contrainte (Burk, 2007; Eckersley *et al.*, 2003). D'une part, elles doivent permettre la circulation libre et aisée des données en opposant le minimum de critères de propriété intellectuelle. D'autre part, elles sont contraintes de protéger leurs données d'une tentative d'appropriation exclusive par un acteur privé en haussant les protections sur la propriété intellectuelle. Pour sortir de cette impasse, le modèle du logiciel libre et ses licences de type *Copyleft* sont mis de l'avant comme une manière alternative d'assurer à la fois le partage et la protection des données (Eckersley *et al.*, 2003). Cependant, Burk (2007) soutient que les licences et les pratiques du logiciel libre ne sont pas tout à fait adaptées au monde de la recherche scientifique, car leurs structures sociales sont radicalement différentes. Entre autres, Burk souligne que les institutions scientifiques peuvent être vieilles de plusieurs siècles, qu'elles sont profondément intégrées dans les rouages de la société et qu'elles disposent de leurs propres cultures organisationnelles.

#### 1.7.3 Les fonctions des données en science

D'autres auteurs se sont penchés sur la nature et les rôles joués par les données elles-mêmes. Ces écrits ont pour point commun de tenter d'ouvrir la boîte noire des données

scientifiques et d'en révéler les processus sociaux sous-jacents. Dans cette perspective, Birnholtz et Bietz (2003) font ressortir les fonctions à la fois scientifiques et sociales que revêtent les données et Bowker (2000) montre qu'elles portent en elles les enjeux techniques, politiques et éthiques inhérents à leur collecte et leur conservation. D'autres auteurs prennent une approche latourienne pour détailler comment la mise en banque des données est le produit d'une série de négociations sociotechniques (Coopmans, 2006; Pontille *et al.*, 2006).

#### 1.7.4 Les usages des bases de données

Pour mieux saisir le rôle des bases de données en science, plusieurs auteurs ont adopté une méthodologie qualitative près de l'ethnographie (Beaulieu, 2003; Borgman, Wallis et Enyedy, 2007; Brown, 2003; Hine, 2006; Star et Ruhleder, 1994, 1996; Zimmerman, 2007, 2008). Ils ont en commun le désir de comprendre les enjeux et les usages des infrastructures de collaboration dans les pratiques scientifiques quotidiennes. Si les auteurs sont globalement intéressés par les mêmes problématiques, les angles d'analyse et les conclusions varient considérablement. Comme seul consensus, il ressort que les usages des bases de données sont variés, complexes et enracinés dans leur contexte social. Bien qu'importants, les facteurs technologiques n'expliquent pas tout.

Dans le même esprit, d'autres chercheurs ont préféré interroger les transformations épistémologiques qui s'opèrent avec la venue des bases de données dans les laboratoires. Loin d'être un simple support aux activités quotidiennes, les bases de données amènent les scientifiques à définir de nouveaux postulats et de nouveaux critères d'objectivité (Beaulieu, 2001; Fujimura, 1999). Les disciplines s'hybrident de plus en plus avec les mathématiques et l'informatique pour développer des logiciels capables d'analyser la quantité colossale d'information présente sur les bases de données (Lenoir, 1998). Aussi, ce mélange des cultures scientifiques amène les chercheurs à expliciter leurs connaissances tacites qu'ils utilisent au quotidien.

Finalement, d'autres auteurs ont choisi de s'intéresser aux enjeux et aux mécanismes organisationnels qui émergent autour de ces dispositifs techniques. Les objets d'analyse prioritaires par ces travaux sont les normes, règlements, dispositions, sanctions et récompenses qui émergent des usages des dispositifs techniques. Ces arrangements, nommés « régimes » (Bohlin, 2004; Hilgartner, 1995; Milanovic, Pontille et Cambon-Thomsen, 2007) ou « règles de bonne conduite » (Cassier, 1998), évoquent une série de conventions, de routines et de pratiques formelles ou

tacites qui encadrent les échanges de données et équilibrent les intérêts antagonistes qui ont cours dans la collaboration scientifique.

### 1.8 Les questions de recherche

Même s'il est possible d'effectuer des rapprochements entre les différents travaux sur les bases de données, il demeure que la recherche dans ce domaine est très dispersée. Si riche les analyses soient-elles, l'étude des bases de données reste orpheline d'une perspective théorique structurante. Certes, les travaux fondateurs de Merton (1974) et ceux de Latour (1987, 1990, 1999; Latour et Woolgar, 1979) sur le rôle des données en science sont régulièrement cités, mais sans être systématiquement utilisés pour l'analyse des bases de données (à l'exception de Coopmans, 2006).

Notre démarche de recherche naît d'une question transversale, qui émerge de l'extrême diversité des travaux sur les usages des bases de données :

- Comment rendre compte de la pluralité et de la complexité des usages des bases de données scientifiques?

Pour tenter de répondre à cette question, nous sollicitons la notion des *régimes d'échange* développée par Milanovic, Pontille et Cambon-Thomsen (2007) pour décrire et comprendre les pratiques d'échange qui ont cours dans les biobanques<sup>4</sup>. Selon la théorisation des auteurs, un régime d'échange « qualifie les liens entre les parties prenantes et définit les formes (matérielles, relationnelles et juridiques) de l'échange de manière spécifique » (2007: 19).

Nous mobilisons cette théorie pour explorer la diversité des pratiques d'échange de données au sein du réseau de recherche *ArcticNet*. *ArcticNet* coordonne, subventionne et encadre environ 500 chercheurs, étudiants et professionnels œuvrant dans la recherche sur l'Arctique canadien.

---

<sup>4</sup> Les biobanques sont des « collections d'échantillons biologiques d'origine humaine [...] [qui] sont généralement associés à des informations (personnelles et biomédicales) structurées la plupart du temps dans une base de données informatique » (Pontille *et al.*, 2006: 1)

Notre question de recherche spécifique est donc la suivante :

- De quelles natures sont les régimes d'échange des bases de données scientifiques dans *ArcticNef*?

Cette même question peut être développée en quelques sous-questions :

- Comment les régimes d'échange émergent, évoluent et se stabilisent-ils?
- Dans quels contextes ces régimes s'inscrivent-ils?
- Comment les aspects sociaux et organisationnels interviennent-ils dans les régimes d'échange?

On remarque que de telles questions portent en elles un postulat fort : celui de l'existence de régimes d'échange. Or, notre intention n'est pas de prouver l'existence des régimes d'échange, mais d'explorer le potentiel de la notion de « régime d'échange » en tant que cadrage pertinent pour la compréhension des pratiques d'échange et de collaboration autour des bases de données scientifiques.

Nous reviendrons plus en détail sur la théorie des régimes d'échange dans le cadre théorique du présent travail, mais soulignons brièvement que cet angle de recherche présente l'avantage de lier plusieurs travaux théoriques et empiriques autour de la notion de « régime » d'une part (Auray, 2004, 2006; Bohlin, 2004; Boltanski, 1990, 1993; Goldenberg, 2010; Hilgartner, 1995; Milanovic, Pontille et Cambon-Thomsen, 2007; Ogien et Quéré, 2005b) et de l'idée de l'émergence de règles dans les communautés scientifiques ou médiatisées par ordinateur d'autre part (Baym, 1995; Cassier, 1998).

### 1.9 La pertinence communicationnelle

Notre programme de recherche s'inscrit dans le champ des études en communication à la fois par son objet d'étude et par les traditions théoriques qu'il mobilise. Dans les prochaines lignes, nous exposons en quoi notre projet peut apporter une contribution appréciable aux études en communication.

Premièrement, notre démarche emprunte grandement à l'étude des usages des technologies de communication (Breton et Proulx, 2002; Jouët, 2000). Au niveau théorique, notre recherche poursuit l'effort entrepris par les études sur les usages pour démontrer comment la technique et le social se construisent mutuellement. Au niveau méthodologique, cela nous amène à opter pour une méthodologie qualitative « [...] pour observer le plus finement possible l'action effective de la technique dans la société à travers une description des usages des objets techniques qu'en font les gens » (Breton et Proulx, 2002: 254).

Deuxièmement, notre étude répond à l'appel de Jones *et al.* (2004) qui demande aux chercheurs en communication de mieux étudier les dynamiques à l'œuvre entre les questions organisationnelles et technologiques. Pour ces auteurs, le développement des TIC est concomitant avec l'évolution des formes organisationnelles. Le travail des chercheurs en communication est donc de mettre en évidence les liens entre les nouvelles technologies et les nouvelles formes organisationnelles. Notre recherche s'inscrit dans cette lignée par son examen des aspects sociaux et organisationnels des bases de données.

Troisièmement, nous espérons que notre recherche parvienne à mettre en relation les champs de recherche en communication et en science, technologie et société (STS). En effet, notre objet de recherche se situe aux frontières des deux champs; les bases de données soulèvent à la fois des questions communicationnelles, technologiques et scientifiques. Comme Boczkowski et Lievrouw (2007a) l'ont observé, les STS et la communication se rejoignent en de nombreux points. Dans un texte plus ancien, Star (1988: 197) décrivait la sociologie de la science et de la technologie dans des termes qui rappellent la communication : « [un champ d'études] petit, vivant [*lively*] et rempli de débats, d'hybridations et [...] de collaborations surprenantes ». Comme nous le verrons dans le chapitre consacré au cadre théorique, notre recherche se propose de puiser dans les traditions épistémologiques et théoriques de ces deux champs de recherche pour aborder les phénomènes du partage de données en science. Ainsi, notre mémoire entend profiter des riches échanges qui ont lieu entre ces deux champs de recherche et contribuer à leur rapprochement.

Finalement, l'inscription de notre démarche dans la communication nous invite à lier notre problématique aux questions fondamentales qui traversent ce champ. Ainsi, nous souhaitons profiter de cette recherche pour développer plus en profondeur des interrogations sur la place de la communication dans les sciences sociales, la distinction entre les communications humaines et



## CHAPITRE II

### CADRE THÉORIQUE

Dans les prochaines pages, nous voyons comment notre question de recherche et notre problématique peuvent être appréhendées par un ensemble de perspectives et de travaux théoriques. Pour ce faire, nous exposons nos choix théoriques et métathéoriques en partant des orientations plus générales pour en arriver à un cadre conceptuel précis.

Nous commençons par lever le voile sur les fondements, sur l'*épistémé* de nos savoirs en clarifiant les positions métathéoriques que nous adoptons (Dortier, 2004). Ensuite, nous passons en revue trois ensembles de travaux qui alimentent notre cadre théorique, soit la « nouvelle » sociologie des sciences et des techniques, l'informatique sociale et la sociologie des usages. Enfin, nous exposons la théorie des régimes d'échange et justifions la pertinence de celle-ci en regard de notre problématique.

#### 2.1 Les questions métathéoriques

Même abstraites, les théories ne se développent pas d'elles-mêmes. Il est donc important de cerner les bases philosophiques sur lesquelles les théories sont bâties. On doit donc recourir à la métathéorie, soit aux théories sur les théories. En reprenant les propos de Miller (2005) et de Burrell et Morgan (1979) sur les questions métathéoriques, nous examinons plus en détail les questions d'ordre ontologiques, épistémologiques et axiologiques.

##### 2.1.1 L'ontologie

L'ontologie porte son attention sur la nature de l'être. Elle s'intéresse à « la nature des phénomènes étudiés, sur le "quoi" de la théorisation » (Miller, 2005: 26). Si plusieurs positions ontologiques existent, nous ne retenons ici que celles proposées par Burrell et Morgan (1979) et

Miller (2005). La première position est la position réaliste. Selon cette position, « le monde social est externe à la cognition humaine, il est tangible, réel et solide » (Burrell et Morgan, 1979: 4). Le monde physique et social est indépendant de la perception. La deuxième position est la position nominaliste, laquelle stipule « que le monde social externe à un individu n'est rien de plus que des noms, des concepts et des étiquettes pour structurer la réalité » (Burrell et Morgan, 1979: 4). Pour un nominaliste, il n'existe pas de monde hors de la perception et de la cognition humaine. La dernière position ontologique est celle du constructivisme social. Selon ce point de vue, le monde social n'est ni totalement objectif, ni totalement subjectif. La réalité est plutôt « une construction intersubjective qui se bâtit à travers les interactions » (Miller, 2005: 27). La réalité sociale n'est pas antérieure à l'action humaine, mais son produit (Leeds-Hurwitz, 1992).

La posture du constructivisme social est fort probablement la plus appropriée pour notre recherche. Selon cette perspective, les technologies et les structures sociales sont mutuellement construites par leurs actions respectives. Ainsi, notre objet d'étude n'est pas vu comme une entité définie, stable et réelle, mais comme le résultat d'interactions variées impliquant les humains et les systèmes techniques. En stipulant que la réalité sociale est issue des interactions, la posture constructiviste amène donc à « rediriger notre attention des produits vers les processus » (Jackson, Poole et Kuhn, 2002: 238). Dans cet esprit, nous sommes amenés à documenter les processus sociaux et communicationnels par lesquels les objets techniques sont appropriés, modifiés, assimilés et intégrés dans la vie de leurs usagers. La communication devient une dimension centrale dans le cadrage constructiviste : c'est par elle que se construit le monde social et c'est par elle que les rapports entre la technique et le social sont négociés.

### 2.1.2 L'épistémologie

L'épistémologie s'intéresse à la création et à l'accumulation des connaissances. Comme pour l'ontologie, il existe plusieurs postures épistémologiques. L'objectivisme est une position épistémologique qui affirme qu'il est possible de comprendre et d'expliquer le monde social par la méthode scientifique, impliquant, entre autres, une séparation entre le chercheur et son objet de recherche (Miller, 2005: 28-29). À l'inverse, la position subjectiviste soutient que « le monde social est relatif à chacun et qu'il convient donc de le comprendre du point de vue de ceux qui sont directement impliqués dans celui-ci » (Burrell et Morgan, 1979: 5).



On ne peut jumeler n'importe quelle position ontologique avec n'importe quelle position épistémologique. Ainsi, une posture ontologique réaliste s'accorde plus facilement avec une posture épistémologique objectiviste et les ontologies nominalistes et constructivistes conviennent mieux aux positions épistémologiques subjectivistes (Miller, 2005).

Pour obéir aux impératifs de cohérence que nous venons d'énoncer, notre recherche doit s'appuyer sur une épistémologie subjectiviste. Tout en étant plus cohérente avec notre ontologie constructiviste (Miller, 2005), la posture subjectiviste s'accorde mieux au type de démarche que nous envisageons et aux connaissances que nous désirons générer. En effet, nous cherchons à comprendre comment les usagers perçoivent, comprennent et s'approprient les bases de données. Cette volonté répond bien à l'hypothèse subjectiviste selon laquelle la réalité doit être comprise du point de vue de ceux et celles qui la vivent. Nous œuvrons donc pour une compréhension plus fine et plus détaillée des usages et de leurs contextes.

### 2.1.3 L'axiologie

L'axiologie étudie les questions relatives aux valeurs dans la recherche (Miller, 2005: 30-31). Selon la formule-choc d'Anderson et Baym, il s'agirait du « théâtre sanglant du politique dans la vie savante » (2004: 609). Selon Miller, le point de vue classique en science considère que les valeurs ne devraient jouer aucun rôle dans la recherche, mais cette posture est nuancée par plusieurs qui ne croient pas que les valeurs peuvent être complètement éliminées des processus de recherche. Une première école de pensée stipule que les valeurs ne peuvent pas être évacuées des processus de définition des problèmes, mais que celles-ci doivent être éliminées du processus de vérification des hypothèses. *A contrario*, d'autres affirment qu'il est impossible d'éliminer les valeurs de quelque partie du processus de recherche que ce soit, car certaines valeurs sont tellement imbriquées dans nos modes de pensée que nous y recourons inconsciemment (Phillips, 1992: 142). Finalement, une troisième position soutient que la recherche elle-même prend part au changement social et devient un acte politique (Guba, 1990: 24).

En matière d'axiologie, il nous semble que l'option la plus honnête est d'affirmer que nous ne pouvons pas éliminer les valeurs du processus de recherche dans son entier. Ainsi, une introspection rapide nous révèle que notre recherche est connotée entre autres par notre volonté d'améliorer la collaboration et le partage des données entre les chercheurs. Si nous avons choisi

d'investir plusieurs mois dans la réalisation d'une recherche et d'un mémoire sur cette question, c'est que nous y trouvions un sens et une motivation.

Notre position de chercheur en communication n'est pas très éloignée de celle des chercheurs du réseau *ArcticNet* que nous avons choisi d'étudier. Les enjeux et les problématiques vécues par ces gens sont en grande partie similaires aux nôtres. En accord avec nos orientations ontologiques et épistémologiques, nous reconnaissons d'emblée le rôle de notre subjectivité dans toutes les facettes de notre recherche. En étant sensible à l'impact des valeurs sur notre démarche, nous nous devons donc d'être plus attentif à nos propres inférences.

## 2.2 Les grandes perspectives théoriques sur la science et la technologie

Maintenant que nous avons précisé nos ancrages ontologiques et épistémologiques, nous discutons brièvement des traditions théoriques que nous souhaitons mobiliser pour notre recherche.

### 2.2.1 La « nouvelle » sociologie des sciences et des technologies

Au courant des années 1970 et 1980, plusieurs chercheurs en sciences sociales ont entrepris de renouveler les approches développées jusque-là dans les études sur les sciences et les technologies (Star, 1988). Jusqu'alors, la compréhension de ces phénomènes était dominée par les approches fonctionnalistes inspirées des travaux de Robert K. Merton ou par les études en bibliométrie. En opposition aux canons mertonniens, une « nouvelle » sociologie des sciences et des technologies (NSST) s'est construite à partir d'idées radicalement différentes des études positivistes. Nous retenons trois idées majeures de ces travaux.

Premièrement, la NSST prend pour hypothèse que l'étude des sciences et des technologies permet de répondre à des « questions centrales à toute la sociologie » (Callon, 1994; Star, 1988: 197). Plus qu'une sous-branche d'une discipline plus vaste, l'étude des phénomènes techniques et scientifiques agit comme un prisme à partir duquel le chercheur peut analyser des problématiques sociales universelles. À titre d'exemple, Michel Callon (1994) relate le travail à la fois technique et social d'ingénieurs travaillant sur un prototype de voiture électrique. En plus de devoir résoudre les problèmes d'ingénierie qu'on leur confie, ces ingénieurs en viennent à construire une société hypothétique dans laquelle cette voiture roulerait. Pour Callon, les arguments et les hypothèses développés par ces ingénieurs les transportent dans le champ de l'analyse sociologique, ce qui fait

d'eux des « ingénieurs-sociologues ». Ainsi, Callon explique que l'analyse des techniques constitue une manière privilégiée d'observer en direct la formation et le développement d'hypothèses sociologiques.

Deuxièmement, la NSST s'appuie sur un cadre épistémologique antipositiviste et met de l'avant la construction sociale de la science (Latour et Woolgar, 1979) et de la technologie (Bijker, Hughes et Pinch, 1994; Jackson, Poole et Kuhn, 2002). Cette perspective constructiviste amène les chercheurs à affirmer la contingence sociale des sciences et des technologies et d'en examiner les processus de construction. À cet effet, l'ouvrage d'ethnographie de laboratoire de Latour et Woolgar (1979) montre comment la production de faits scientifiques est guidée par une série intense d'actions, de négociations et d'ajustements. Sans être fallacieux, ces faits « expérimentaux » sont davantage tributaires des conditions sociales qui prévalent au sein d'un laboratoire que d'une méthode scientifique universelle et objective.

Finalement, les chercheurs de la NSST se tournent vers des méthodologies ethnographiques (p. ex. Fujimura, 1999; Latour et Woolgar, 1979) ou historiques (p. ex. Hughes, 1994; Lenoir, 1998) dans leurs travaux. Pour Star (Star, 1988: 198), ces méthodologies sont des voies royales pour ouvrir la « boîte noire » des sciences et des technologies : en s'imprégnant de la culture et du langage des acteurs, on peut parvenir à reconstituer les processus de construction et de stabilisation des savoirs et des techniques. C'est par la profondeur de l'analyse et la richesse des observations des approches qualitatives que les chercheurs de la NSST peuvent parvenir à la description détaillée de la construction sociale des techniques et des savoirs savants.

### 2.2.2 Les études en informatique sociale

L'informatique sociale (*Social Informatics*) s'est constituée au milieu des années 1990 par la volonté de chercheurs provenant de nombreuses disciplines de rassembler leurs travaux sur les aspects sociaux et culturels des technologies de l'information (Kling, Rosenbaum et Sawyer, 2005: 5-6). L'informatique sociale se veut donc « l'étude interdisciplinaire du design, des usages et des conséquences des technologies de l'information en prenant en compte leurs interactions avec les contextes institutionnels et culturels » (Kling, 1999). L'informatique sociale a développé un cadre théorique qui met de l'avant les interrelations entre les systèmes informatiques et les contextes institutionnels. La perspective de l'informatique sociale nous amène à considérer

attentivement le contexte organisationnel des technologies et la façon dont il intervient dans l'appropriation et les usages d'un dispositif.

Une des idées centrales des travaux en informatique sociale est que les « TIC *permettent et contraignent* l'action sociale [nous accentuons] » (Kling, Rosenbaum et Sawyer, 2005: 20). Les documents électroniques en constituent un bon exemple. Par rapport aux documents papier, ils sont facilement accessibles à partir d'un ordinateur connecté à Internet. Par contre, la lecture de ces documents reste limitée aux postes informatiques alors que les copies papier peuvent être transportées, lues et annotées sans difficulté (Idem). Ainsi, tout en ouvrant de nouvelles possibilités, les TIC « restructurent les traitements de l'information déjà existants et les relations sociales qui les entourent » (Kling, Rosenbaum et Sawyer, 2005: 20; Sawyer *et al.*, 2003).

L'informatique sociale présente aussi l'avantage d'avoir développé un pan important de ses recherches sur les communications scientifiques. Rob Kling, une des figures importantes de ce domaine d'étude, s'est intéressé de très près aux communications scientifiques à l'ère d'Internet et de la communication médiatisée par ordinateur (Kling, 2004; Kling et Callahan, 2003; Kling et Crawford, 1999; Kling et McKim, 1999; Kling, McKim et King, 2003).

Nous retenons de l'informatique sociale l'importance de considérer le contexte plus large dans lequel s'insère une technologie donnée. Par exemple, Kling (1999) a montré comment les structures de récompenses et de sanctions d'une firme comptable ont rendu son système de collaboration informatisée complètement inutile. De la même manière, nous souhaitons comprendre comment les contextes organisationnels et scientifiques influent sur la forme des usages des bases de données en science.

### 2.2.3 La sociologie des usages

Un peu à la manière de l'informatique sociale et de la NSST, les études en sociologie des usages se sont développées à partir des années 1980 pour « [mettre] au jour le rôle actif de l'utilisateur dans le modelage des emplois de la technique » (Jouët, 2000: 493). Nourrie par l'apport de Michel de Certeau (1990) sur les détournements des usages et unie dans le rejet du déterminisme technique, la sociologie des usages prend le parti d'« [...] observer le plus finement possible l'action effective de la technique dans la société à travers une description des usages des objets techniques qu'en font les gens » (Breton et Proulx, 2002: 254). Cette approche milite pour

une méthodologie qualitative près de l'ethnographie, pour être en mesure de rendre compte de la profondeur sociale des usages.

Même si elle fait l'objet de critiques en raison de son trop grand empirisme, l'éclatement de ses objets de recherche (Jouët, 2000) ou son vide critique (Granjon, 2004), la sociologie des usages nous offre un cadre d'analyse très pertinent pour comprendre comment l'on s'approprie les TIC. Plutôt que d'être inscrits à même les caractéristiques des objets techniques, les usages y sont conceptualisés dans une logique d'émergence et d'évolution : « Faut-il rappeler que le développement de nouveaux usages n'émerge pas ex nihilo [...] ? L'apparition de nouvelles pratiques se greffe sur le passé, sur des routines, sur des survivances culturelles [...] » (Mallein et Toussaint, 1994: 317). Cette perspective de « généalogie des usages » (Jouët, 2000) nous amène à interroger l'ancrage historique des technologies (p. ex. Flichy, 1991) et à explorer les conditions et les phases par lesquelles leurs usages intègrent le quotidien.

#### 2.2.3.1 L'utilisateur contributeur au cœur du Web social

Les études sur les usages peuvent également être liées avec les conceptualisations plus récentes sur le rôle des usagers dans le développement des dispositifs technologiques. Depuis une dizaine d'années, des études issues de la sociologie des usages tentent de renverser le modèle binaire concepteur-consommateur en montrant comment les usagers sont de plus en plus appelés à participer au développement des technologies.

Beaucoup ont porté leur attention sur les multiples avatars du Web social (Millerand, Proulx et Rueff, 2010). En effet, il appert que l'utilisateur du Web social quitte son rôle passif de consommateur en générant des contenus et en les partageant. La différence entre concepteur et utilisateur devient alors très ténue. Comme exemple de cette logique, *Google* met en place le site *Google Code* qui fournit la documentation et les références nécessaires pour permettre aux développeurs de modifier et de personnaliser les applications de l'entreprise.

Dans le Web social, la figure de l'utilisateur est donc appelée à être graduellement remplacée par celle de l'utilisateur-contributeur, lequel est à la fois producteur et consommateur de contenus et d'innovation (Proulx, 2010). Cette nouvelle conceptualisation de l'utilisateur s'actualise dans les projets d'infrastructure scientifique, dans lesquels le partage et la contribution de données sont attendus de tous. Comme mentionnés précédemment, les développements du Web social ont

inspiré les promoteurs des laboratoires et des bases de données scientifiques (Burk, 2007; Eckersley *et al.*, 2003).

#### 2.2.3.2 Usages et rapports sociaux

Notre réflexion critique sur les bases de données scientifiques nous vient de l'articulation des usages avec les rapports sociaux. Comme l'ont démontré Mallein, Toussaint et Bydlowski (1984) pour le Télétel et Lelong, Thomas et Ziemlicki (2004) pour l'Internet, l'introduction d'un nouveau dispositif technologique dans les foyers suit un parcours fortement déterminé par les rôles sexués : les femmes ont très peu de temps pour se familiariser avec le nouvel appareil, puisqu'elles sont encore largement responsables des travaux domestiques. Jouissant généralement de plus de temps discrétionnaire, les hommes s'approprient plus rapidement les outils au moment même où ces derniers deviennent des objets « tout puissants ». De même, Jouët (2000: 508) observe que des logiques similaires ont lieu dans les entreprises où les employés les moins qualifiés perdent une part de leur autonomie professionnelle avec l'introduction de nouveaux systèmes informatiques. Ces travaux montrent bien comment les usages technologiques restent enchâssés dans des structures et des formes de dominations sociales plus fortes (patriarcat, capitalisme, autoritarisme, etc.).

En reprenant à notre compte les questionnements critiques suscités par la sociologie des usages, nous sommes amenés à considérer notre objet de recherche sous un nouvel angle. Au-delà des bénéfices promis par les bases de données, il a été démontré que de tels projets pourraient représenter un fardeau énorme en terme de travail et d'effort. Les promoteurs de projets doivent mettre en place l'infrastructure technique, veiller au financement, aux communications, à l'entretien et au développement continu de leur projet (Keil, 1990; Star et Ruhleder, 1994, 1996). Les contributeurs se doivent, quant à eux, d'apprendre à se servir des outils du système, préparer leurs données et les soumettre, tout en restant en communication avec les gestionnaires de la base de données (Beaulieu, 2003). Il est à craindre que cette charge de travail soit très inégalement répartie entre les acteurs du monde de la recherche. Souvent peu valorisés, ces tâches ont de bonnes chances d'échoir aux assistants de recherches, aux étudiants ou aux techniciens de laboratoire, bref, aux acteurs situés en bas de la hiérarchie universitaire. En se fiant aux travaux antérieurs sur les usages, il est à prévoir que les rapports sociaux présents dans les laboratoires se traduiront par des pratiques différenciées en matière d'usages des bases de données.

Cette piste d'analyse que nous propose la sociologie des usages nous permet de lier nos intérêts de recherche avec plusieurs thématiques porteuses, dont notamment celle du travail invisible (Shapin, 1989; Star et Strauss, 1999).

### 2.2.3.3 Usages et travail invisible

À la fin des années 1980, Steven Shapin (1989) a révélé une problématique majeure des études en histoire et en sociologie des sciences : le travail effectué par les techniciens est systématiquement rendu invisible. En analysant les archives du chimiste Robert Boyle (1627-1691), Shapin met au jour comment de nombreux manœuvres, techniciens et assistant anonymes ont effectué l'essentiel du travail physique, technique et intellectuel derrière l'œuvre de Boyle. Si cette situation peut être attribuée aux rapports sociaux du 17<sup>e</sup> siècle, Shapin remarque que les historiens et les sociologues des sciences continuent d'occulter le travail des techniciens dans leurs travaux.

À la suite de la publication de son article sur le *technicien invisible*, la pensée de Shapin a trouvé de nombreux échos (p. ex. Dietrich et Tambasco, 2007). Entre autres, Star et Strauss (1999) ont repris la thématique du travail invisible pour montrer comment celle-ci s'actualise dans le travail coopératif assisté par ordinateur (*Computer Supported Cooperative Work*, CSCW). Les auteurs vont plus loin en proposant un modèle de « mise en invisibilité » qui catégorise les processus par lesquels les travailleurs et leur travail deviennent invisibles. Par la *création d'une non-personne*, *disparition du travail dans le décor* et *l'abstraction et la manipulation d'indicateurs*, Star et Strauss soutiennent que la *mise en invisibilité* du travail peut prendre plusieurs formes. Par ces processus, la contribution de certains usagers peut être réduite à l'invisibilité dans des projets tels que des bases de données scientifiques. Par exemple, des travaux récents ont montré comment les techniciens de ces systèmes réalisent un travail important de gestion et d'entretien des jeux de données. Pourtant, les discours d'accompagnement des bases de données scientifiques que nous avons recensés au dernier chapitre (Arzberger *et al.*, 2004b, 2004a; Drake, 2003; National Research Council, 1997; Schroder, 2003; Wellcome Trust, 2003) restent muets sur le rôle et la contribution des techniciens et des gestionnaires de données dans les projets d'infrastructures scientifiques.

### 2.3 Une piste pour saisir les modalités d'usage : les régimes d'échange

Le terme régime évoque, tant en français qu'en anglais, un système organisé de règles (Dictionary of the Social Sciences, 2002). Il est employé abondamment en science politique (régime parlementaire, ancien régime, etc.). De manière générale, « régime peut aussi signifier un cadre large de règles et de normes qui administrent un cas particulier » (Idem). L'idée de régime s'est graduellement imposée dans notre programme de recherche à mesure que nous développons une problématique et rassemblions la littérature pertinente sur les bases de données scientifiques.

#### 2.3.1 Les régimes d'action

Une première vague de théorisation a émergé à partir de la fin des années 1980, alors qu'une poignée de chercheurs intéressés aux théories de l'action (Ogien et Quéré, 2005a; Quéré, 1993) ont mobilisé la notion de régime. Luc Boltanski et Laurent Thévenot (1991) proposent une « sociologie dite des "régimes d'action" [qui] analyse les formes de la coordination des actions et explore les appuis conventionnels, les ressources et les capacités dont les agents disposent pour s'ajuster entre eux et à leur environnement » (Ogien et Quéré, 2005b: 105). Boltanski (1990) poursuit son raisonnement et propose de distinguer les régimes *sous équivalences* et *hors équivalences* et des régimes *de paix* et *de disputes*. De ces distinctions, il propose un schéma de quatre régimes d'actions distincts et incommensurables (paix hors équivalence, paix sous équivalence, dispute hors équivalence, dispute sous équivalence). Chaque régime organise et coordonne les actions des acteurs qui tendent à se maintenir durablement dans celui-ci. L'intérêt de Boltanski se situe donc dans les moments où les acteurs vainquent l'inertie et passent d'un régime à un autre.

Ce modèle des régimes d'action développé par Boltanski et Thévenot a un certain succès au sein de la sociologie française. Corcuff (1998) y voit un modèle alternatif aux « sociologies bulldozer » de Bourdieu, Crozier, etc., qui aplatissent les aspérités du social au moyen de concepts totalisants comme l'habitus, le capital ou le pouvoir. Pour Corcuff, il s'agirait d'un modèle suffisamment heuristique et souple pour mettre au jour les dynamiques d'action et de justification vécues en société. Notamment, le modèle des régimes d'action a été utilisé par Corcuff pour examiner la question de la compassion dans deux *Caisses d'Allocations Familiales* et pour reconstruire les processus de décision sur la construction de l'aéroport Lyon-Saint Exupéry (Corcuff, 1996, 1998).



### 2.3.2 Les régimes pour comprendre les communautés du logiciel libre et les *wikis*

Plus près de nos intérêts de recherche, Nicolas Auray (2004, 2006) a repris l'apport de Boltanski et Thévenot en développant l'idée des *régimes d'engagement* pour explorer les problématiques vécues dans les communautés de développeurs du logiciel libre. Au terme de son étude, Auray conclut que les régimes d'engagement qu'il a observés concilient des forces antagonistes : « [...] Ils articulent une participation fondée sur le don et un lien d'intérêt marchand. Ils font vivre ensemble des groupes d'utilisateurs aux normes et intérêts dissemblables » (2006: 19).

Dans un texte récent tiré de l'ouvrage *Web social*, Goldenberg (2010) discute de la portée sociale du don au sein des collectifs de développeurs de logiciel libre et les *wikis*. Afin de distinguer don et contribution au sein de ces groupes, Goldenberg mobilise le modèle des régimes d'action de Boltanski (1990; voir aussi Nachi, 2003) pour examiner en quoi le don s'inscrit davantage dans les régimes *hors équivalences* qui « [...] renvoie[nt] à un état fusionnel ainsi qu'à la gratuité, la générosité ou la bienveillance qui se passe de calcul ou de mise en équivalence. » (Goldenberg, 2010: 226). Quant à elles, les contributions se placent dans les régimes *sous équivalences* dans lesquels « [...] la justesse [de la contribution] est évaluée en fonction d'un intérêt collectif. En cas de désaccord, le contributeur doit justifier son propos, au risque de voir retirer les éléments non justifiés de sa participation » (2010: 227).

Nous n'avons présenté qu'une fraction minuscule de la pensée des chercheurs actionnistes autour de l'idée de régime et n'avons abordé que superficiellement leurs échos contemporains. Il s'agit simplement d'une toile de fond sur laquelle nous développons davantage la notion de régime appliquée au cas de la communication scientifique.

### 2.3.3 Les régimes de communication scientifique

Dans un texte fréquemment cité, Hilgartner (1995) utilise les positions théoriques de la NSST pour poser la notion de régimes communicationnels pour expliquer les formes de communication scientifique permises par les bases de données. En documentant les progrès rapides des bases de données en biologie, Hilgartner émet l'hypothèse que l'on assisterait à l'émergence d'un nouveau régime communicationnel en science.

Pour Hilgartner, un régime est un « système sociotechnique socialement et historiquement contingent de communication des résultats des recherches [...] *contraignant et permettant* l'action des usagers [nous accentuons] » (Bohlin, 2004: 384; Hilgartner, 1995: 244-245). À partir de cette définition, Hilgartner oppose l'« ancien » régime des revues scientifiques au « nouveau » régime des bases de données. Alors que le premier est relativement institutionnalisé, le second amène une série de changements et de redéfinitions : les frontières entre le publié et le non-publié se brouillent, les mécanismes de récompenses restent à définir, etc.

Si les travaux d'Hilgartner ont fait l'objet de critiques, notamment pour avoir surestimé l'impact des bases de données dans les pratiques scientifiques (Hine, 2006), la notion de régime a été retenue pour l'analyse des bases de données scientifiques. Entre autres, Bohlin (2004) l'a reprise pour interroger les relations entre les revues scientifiques officielles et les initiatives de publications scientifiques libres sur Internet (*Open Access*). Plus récemment, la notion de régime a été réactualisée par Milanovic, Pontille et Cambon-Thomsen (2007) pour rendre compte des pratiques d'échange dans les biobanques.

#### 2.3.4 Les régimes d'échange

La démarche de Milanovic *et al.* (2007) part du constat que le schéma binaire *source de données-utilisateurs de données* est dominant dans notre compréhension des pratiques d'échange dans les biobanques. *A contrario*, ces échanges feraient intervenir un grand nombre d'acteurs hétérogènes et selon des modalités très complexes. Par exemple, certains chercheurs s'échangent des données de manière très informelle alors que d'autres accompagnent les échanges de plusieurs dispositions légales et financières. Pour tenter d'expliquer et d'ordonner cette pluralité des échanges, Milanovic, Pontille et Cambon-Thomsen mènent une étude empirique auprès de trois biobanques françaises aux tailles et aux missions différentes. À l'aide d'une méthodologie qualitative et inductive inspirée de la théorie ancrée (Glaser et Strauss, 1967), ils dégagent trois régimes d'échange distinctifs qui cohabitent dans les usages des biobanques :

1. *Le régime de coopération* : ce régime repose sur la confiance et n'est que très peu encadré formellement. Les gens s'échangent échantillons et données de manière implicite et non problématique. Il s'agirait d'habitudes qui se sont développées avec le temps. Les acteurs impliqués dans ces échanges (médecins, chercheurs) se

côtoient naturellement dans leur vie professionnelle. Tous reconnaissent la place et les rôles de chacun.

2. *Le régime de don* : ce régime est basé sur la réciprocité : « je te fais un don (données, analyses, échantillons, avis légaux) en échange d'un service futur ou d'un autre don (citation, être reconnu comme auteur dans les travaux de recherche, etc.) ». En s'inspirant des travaux de Mauss (1950) sur la portée sociale du don, ce régime de don peut être considéré comme une forme réifiée des dynamiques qui ont cours dans ce milieu.
3. *Le régime de sous-traitance* : il s'agit d'un régime où les biobanques agissent comme n'importe quelle entreprise : elles offrent et facturent leurs services. Les échanges sont hautement codifiés, écrits et juridiques. Cependant, Milanovic, Pontille et Cambon-Thomsen observent qu'une bonne part de ces échanges reste dépendante de la confiance mutuelle entre les partenaires.

Pour Milanovic, Pontille et Cambon-Thomsen, ces régimes d'échange ne sont pas mutuellement exclusifs. Les trois régimes peuvent coexister à l'intérieur d'une même biobanque. Les auteurs soutiennent également que de tels régimes remettent en question la distinction dualiste entre la science ouverte conforme aux idéaux de Merton (1974) et la recherche scientifique privée. En effet, des institutions aux objectifs parfois antagonistes parviennent à collaborer autour des biobanques et à établir des relations mutuellement avantageuses (Cassier, 1998).

Ainsi, notre recherche s'alimente de la théorie des régimes d'échange pour enrichir notre compréhension des pratiques de partage des données en science. En cohérence avec nos bases épistémologiques et les courants de recherche que nous avons mobilisés, ce cadre théorique nous sert de tremplin initial pour développer notre analyse des pratiques de partage de données et décrire les contextes dans lesquels elles ont lieu.

#### 2.3.5 La pertinence de la théorie des régimes d'échange

La typologie des régimes d'échange possède, selon nous, une valeur heuristique très forte, dans le sens où elle permet aisément de différencier et de catégoriser les pratiques de partage et de collaboration scientifique qui nous intéressent. De plus, par son recours au concept de régime et

par son emprunt à la sociologie du don (Mauss, 1950), ce cadre conceptuel se lie harmonieusement à nos ancrages théoriques et, par le fait même, ouvre de nouvelles pistes d'analyse.

La théorie des régimes d'échange nous est apparue comme idéale pour ce mémoire parce qu'elle indique quels sont les indices, les indicateurs et les propriétés qui caractérisent chacun des régimes. Ainsi, à la lecture de l'article de Milanovic, Pontille et Cambon-Thomsen, nous sommes invités à considérer très attentivement les questions de confiance, de proximité, de réciprocité et de formalisation des échanges dans le partage des données scientifiques. Comme nous le verrons dans le chapitre sur le cadre méthodologique, ces indicateurs ont été cruciaux dans la réalisation du codage du corpus et de la constitution des catégories d'analyse.

Si elle se révèle remarquablement adaptée à nos besoins, la théorie des régimes d'échange comporte certaines lacunes. D'abord, l'article de Milanovic, Pontille et Cambon-Thomsen (2007) n'a pas été cité ou critiqué dans des revues majeures<sup>5</sup>. Cela suppose donc que nous ne pouvons pas nous référer à d'autres ouvrages et d'autres auteurs pour nous livrer à la discussion et à la critique de cette théorie. Ainsi, nous nous trouvons sur un terrain peu balisé et sur lequel peu d'auteurs se sont penchés. À défaut de pouvoir faire intervenir des critiques externes à la théorie des régimes d'échange, nous pouvons cependant creuser dans les autres travaux des auteurs de la théorie. En l'absence de regards externes sur la théorie, l'examen du cheminement intellectuel qui la précède nous permet de mieux en saisir le contexte et les ancrages.

La théorie des régimes d'échange a pour effet d'éloigner notre attention initialement centrée sur les dispositifs techniques pour l'orienter vers les processus sociaux. Un tel passage est tout à fait en cohérence avec nos bases constructivistes, lesquelles nous invitent à « rediriger notre attention des produits vers les processus » (Jackson, Poole et Kuhn, 2002: 238). Par contre, l'emploi de cette théorie relègue à l'arrière-scène les questions technologiques, lesquelles étaient

---

<sup>5</sup> Malgré sa remarquable clarté, l'effort de théorisation entrepris par les auteurs a eu très peu d'échos. Une recherche sur les principales bases de données bibliographiques ne nous a pas permis de retrouver des textes citant l'article de Milanovic, Pontille et Cambon-Thomsen (2007). Il convient toutefois de rappeler que l'article est récent. Selon nous, cette conceptualisation mérite d'être davantage explorée.

pourtant au cœur des premières interrogations de cette recherche. En effet, la question des dispositifs techniques qui rendent possibles les échanges est peu abordée par les auteurs.

### 2.3.6 Les définitions des concepts utilisés

Pour opérationnaliser notre question en un programme de recherche cohérent, il importe de définir les concepts centraux que nous comptons utiliser. Ainsi, nous explicitons les concepts de base de données scientifique et de régime d'échange.

- En reprenant la définition développée dans la problématique, une *base de données scientifique* est un dispositif informatique matériel ou logiciel connecté à un réseau qui permet le dépôt, l'archivage, l'organisation, l'accès et le traitement d'informations et de données primaires de recherche par un grand nombre d'utilisateurs.
- Un *régime d'échange* « qualifie les liens entre les parties prenantes et définit les formes (matérielles, relationnelles et juridiques) de l'échange de manière spécifique » (Milanovic, Pontille et Cambon-Thomsen, 2007: 19).

Milanovic, Pontille et Cambon-Thomsen proposent trois formes différentes de régimes d'échange : le régime de coopération, de don et de sous-traitance.

- Le *régime de coopération* est un régime d'échange très peu formalisé qui repose sur la familiarité des acteurs et sur un très haut niveau de confiance mutuelle.
- Le *régime de don* est un régime d'échange reposant sur la réciprocité des dons entre acteurs.
- Le *régime de sous-traitance* est un régime d'échange basé sur une offre commerciale de services qui est hautement formalisée (contrats, traces écrites, documents juridiques, etc.).

Dans la mesure où notre recherche vise à comprendre les phénomènes de partage des données à travers une perspective épistémologique interprétative, nous ne nous limitons pas au seul concept des régimes d'échange. D'autres cadres conceptuels peuvent être sollicités en cours de route.

Dans les dernières pages, nous avons fait état de notre cheminement intellectuel qui nous a conduit à mobiliser la théorie des régimes d'échange (Milanovic, Pontille et Cambon-Thomsen, 2007) pour répondre à notre question de recherche. D'abord, nous avons cerné nos ancrages en traitant successivement des postures ontologiques, épistémologiques et axiologiques que nous prenons. Ensuite, nous avons discuté de trois courants de recherche qui s'accordent à la fois à notre problématique et à nos *a priori* métathéoriques. Enfin, nous avons vu en quoi la théorie des régimes d'échange nous offre l'opportunité de combiner ces courants théoriques en un cadre heuristique pour développer notre compréhension des pratiques de partage de données.

Notre cadrage conceptuel étant précisé, le prochain chapitre est consacré à la présentation de notre démarche méthodologique.

-- -- -- -- -- -- -- -- -- --

## CHAPITRE III

### CADRE MÉTHODOLOGIQUE

Dans ses grandes lignes, notre démarche de recherche générale suit le cheminement d'une méthodologie qualitative relativement « classique ». En premier lieu, nous avons procédé à un méticuleux travail de problématisation et de théorisation au courant de l'année universitaire 2008-2009. Au terme de cette démarche, nous sommes parvenus à cerner notre objet de recherche, notre problématique, la documentation scientifique pertinente ainsi qu'un modèle théorique pour l'analyser. Nous avons rapporté les résultats de ces étapes dans les deux chapitres précédents. À l'automne 2009, nous avons procédé à la collecte des données, que nous avons ensuite analysées à l'hiver 2010. Dans les prochaines pages, nous rendons compte de ces deux phases cruciales de notre recherche. Avant d'aller plus loin, nous présentons d'abord le terrain que nous avons retenu pour notre recherche.

#### 3.1 Le terrain de recherche

##### 3.1.1 Une présentation d'*ArcticNet*

Nous avons choisi d'effectuer notre terrain de recherche au réseau de centres d'excellence du Canada *ArcticNet*, lequel rassemble des scientifiques autour de l'étude de l'Arctique. Actif depuis 2003, *ArcticNet* encadre, subventionne et coordonne le travail de recherche dans l'Arctique canadien. Basé à l'Université Laval à Québec, *ArcticNet* regroupe « plus de 145 chercheurs d'ArcticNet ainsi que 580 étudiants des cycles supérieurs, chercheurs postdoctoraux, associés de recherche et techniciens, issus de 30 universités canadiennes, de 8 ministères fédéraux et de 11 départements provinciaux [...] » (ArcticNet, 2010a: 54).

Le réseau est multidisciplinaire, regroupant en son sein les sciences de la nature, les sciences de la santé et les sciences sociales. Il est à noter que les sciences naturelles sont beaucoup

plus représentées au sein d'*ArcticNet* (environ 55 % des projets) que les sciences sociales (32,5 %) ou les sciences de la santé (12,5 %) (*ArcticNet*, 2010b).

*ArcticNet* est un réseau de centres d'excellence (RCE), une classification qui lui permet de recevoir un financement important du *Secrétariat des Réseaux de centres d'excellence du gouvernement du Canada*. En tant que RCE, *ArcticNet* a pour mission de développer des programmes de recherche de haut niveau, mais aussi de maximiser les liens entre l'industrie, les gouvernements et les communautés du Nord. Cela se passe entre autres par la formation de personnel hautement qualifié, le développement de partenariats, le transfert de connaissances et une foule d'initiatives de sensibilisation<sup>6</sup>.

Dans le cadre de son programme scientifique, *ArcticNet* a lancé un projet de développement d'une base de données en ligne de grande envergure, le *Polar Data Catalogue* (PDC). À l'heure actuelle, cette base de données est encore en développement. Elle contient principalement des métadonnées et quelques jeux de données primaires<sup>7</sup>.

Millerand apporte quelques précisions sur le projet du *Polar Data Catalogue* :

— *Plus précisément, le projet vise la constitution d'une des plus importantes bases de données sur l'Arctique au plan international, à travers l'interconnexion de plusieurs systèmes existants et l'intégration des nouvelles données produites par le réseau dans le cadre des projets de recherche de ses membres. Il faut préciser, d'une part que l'archivage des données est au cœur de la mission du réseau ArcticNet et d'autre part, que ce projet d'infrastructure intervient dans un contexte international où la plupart des agences de financement rendent les octrois conditionnels à un « archivage » adéquat des données scientifiques en vue de leur diffusion et, surtout, de leur réutilisation.* (Millerand, 2007: 6)

---

<sup>6</sup> À titre d'exemple, un des programmes de sensibilisation et de transfert de connaissances qu'offre *ArcticNet* est « Écoles à bord », un programme dans le cadre duquel une dizaine d'étudiants du secondaire sont invités à bord du navire de recherche *NGCC Admundsen* pour participer à diverses activités éducatives.

<sup>7</sup> La phase 1 du projet du PDC prévoit la soumission obligatoire des métadonnées de recherche pour les membres du réseau *ArcticNet*. Dans la deuxième phase du projet, il est prévu d'y ajouter les données primaires de recherche.



En bref, le *Polar Data Catalogue* s'inscrit au cœur des objectifs scientifiques et de transfert des connaissances du réseau. Son premier objectif est de constituer une banque de données suffisamment étendue pour ouvrir de nouvelles avenues de recherche et permettre aux chercheurs d'effectuer des analyses croisées sur plusieurs jeux de données. Un deuxième objectif du *Polar Data Catalogue* est de mieux sensibiliser les communautés nordiques à la recherche ayant cours dans leur région en rendant aisément disponibles les métadonnées et les données de recherche. Aussi, le PDC permettra aux fonctionnaires, élus, entreprises et groupes locaux d'accéder à des données importantes pour leurs projets de développement social et économique.

### 3.1.2 La pertinence du terrain

Le choix d'*ArcticNet* comme terrain de recherche s'est fait très aisément, car nous disposons déjà d'un contact privilégié avec le réseau de par notre collaboration à la recherche sur le processus de développement de la base de données d'*ArcticNet* : « Des infrastructures d'information pour les scientifiques : analyse des processus d'innovation sociotechnique au sein d'un projet d'infrastructure de recherche » de notre directrice Florence Millerand (2008-2011, subvention FQRSC). Notre directrice nous a aussi mis en contact avec deux personnes-clés du réseau *ArcticNet*, lesquelles nous ont grandement aidés à rencontrer nos informateurs et à nous familiariser avec la recherche sur l'Arctique.

Au-delà de son accès aisé, le choix d'*ArcticNet* se justifie par sa pertinence théorique. Assurément, *ArcticNet* représente un terrain très riche où il est possible d'observer en direct le développement d'un projet de base de données dans un contexte multidisciplinaire. En effet, nous avons rencontré des chercheurs œuvrant dans les trois grands domaines scientifiques d'*ArcticNet* afin de sonder leurs pratiques et leurs perceptions du partage de données. Telle que mentionnée plus tôt, la compréhension des cultures disciplinaires et des différentes méthodologies est cruciale à l'analyse des pratiques d'échange de données (Carlson et Anderson, 2007; Corti, 2000; Talja et Maula, 2003).

La pertinence du réseau *ArcticNet* en tant que terrain de recherche trouve écho chez Laperrière (1997: 313-314), qui affirme que le choix du terrain doit s'harmoniser à la question de recherche. Conformément aux visées exploratoires et compréhensives de notre recherche, les critères de représentativité ou de d'universalité ne sont pas centraux dans une telle démarche. Notre enquête a plutôt pour objectif d'analyser le cas riche et complexe d'*ArcticNet* pour en faire

ressortir les points saillants et les reliefs plutôt que de chercher des généralités à partir d'un grand nombre de cas.

### 3.2 L'orientation méthodologique

D'emblée, nous abordons notre problématique par l'entremise d'une méthodologie essentiellement qualitative. Notre projet ne cherche ni à prouver, ni à infirmer l'existence de régimes d'échange ou de partage de données, mais plutôt à explorer ces pratiques et à développer notre compréhension de ces phénomènes (Poupart, 1997).

Nous optons pour une démarche qualitative pour deux raisons fondamentales. Premièrement, les outils de cueillette tels que l'entrevue, l'analyse des documents internes et l'observation *in situ* permettent d'observer la subjectivité des acteurs et de rendre compte de la richesse des situations, ce qui concorde avec l'orientation épistémologique de notre recherche. Deuxièmement, l'emploi d'une stratégie de recherche qualitative trouve de nombreux échos dans le cadrage théorique que nous avons mobilisé. En effet, autant la nouvelle sociologie des sciences et des technologies (Star, 1988) que l'étude des usages (Breton et Proulx, 2002; Granjon, 2004; Jouët, 2000; Mallein et Toussaint, 1994) amènent le chercheur à adopter une démarche qualitative.

### 3.3 Les considérations éthiques

Sur le plan éthique, la démarche de la recherche n'a pas différé substantiellement d'autres recherches qualitatives. Naturellement, notre recherche a dû satisfaire les exigences du *Comité institutionnel d'éthique de la recherche avec des êtres humains* de l'Université du Québec à Montréal (UQAM). En juin 2009, la recherche de madame Florence Millerand sur *ArcticNet* a reçu un certificat d'éthique de la part de ce comité. Comme notre projet de mémoire se rattache à cette étude, nous avons reçu notre certification éthique par la même occasion.

Au-delà des démarches institutionnelles obligatoires, nous avons veillé à intégrer nos préoccupations éthiques dans toutes les étapes de notre recherche. Premièrement, notre réflexion a commencé lors de nos lectures préliminaires et de la constitution de la problématique. En lisant en profondeur sur les enjeux du partage de données, sur les contextes scientifiques, sur les infrastructures techniques, etc., nous sommes parvenus à dégager certaines lignes de tension et à cerner certains sujets plus délicats. Par exemple, la question de la compétitivité du milieu

scientifique est revenue très fréquemment dans les textes sur le partage de données. Plusieurs textes ont fait état du malaise qu'exprimaient les scientifiques quant à cette situation et comment cette problématique rendait la coopération difficile. Bref, très tôt dans le développement de notre projet, nous avons été sensibilisés à ces thématiques délicates.

Deuxièmement, lors des entrevues, nous nous sommes assurés que les interviewés comprenaient bien les objectifs et les méthodes de notre projet et qu'ils se sentaient tout à fait à l'aise d'y participer. Avec Madame Millerand, nous avons mis en place un formulaire de consentement relativement succinct qui énonçait les droits des informateurs et fournissait les coordonnées du président du comité éthique de l'UQAM.

Troisièmement, dès que les enregistrements audio des entrevues ont été réalisés, nous avons assigné à chaque informateur un pseudonyme. À partir de ce moment, nous avons utilisé le pseudonyme pour l'analyse et la communication des résultats.

Quatrièmement, nous nous sommes engagés auprès des répondants à leur demander l'autorisation pour toutes citations d'entrevue que nous pourrions utiliser dans nos communications scientifiques. Ainsi, avant de publier un rapport ou de présenter une communication dans un colloque, nous avons contacté chaque informateur cité pour qu'il approuve les citations.

Finalement, nous avons entrepris une réflexion sur l'importance de nos propres valeurs et des rôles qu'elles auraient pu jouer dans notre recherche. Sans aucun doute, notre démarche était motivée par notre désir d'améliorer tangiblement les processus et les outils de collaboration scientifique. Aussi, notre condition de chercheur en communication n'était pas très éloignée de celles des chercheurs que nous souhaitions étudier. En accord avec nos ancrages ontologiques et épistémologiques, nous avons embrassé la subjectivité de notre démarche et avons reconnu que notre savoir était le produit d'un contexte sociétal, organisationnel et interpersonnel. Aussi modeste soit-elle, notre recherche prend part à la construction des connaissances et de nos perceptions de l'activité scientifique, lesquelles ont des conséquences tangibles (politiques publiques sur la recherche, financement, rapports institutionnels, etc.).

### 3.4 La cueillette des données

#### 3.4.1 Les outils de cueillette

Pour la cueillette des données, trois outils ont été mobilisés. Premièrement, nous avons interviewé onze acteurs du réseau *ArcticNet* pour les entendre sur leurs pratiques d'échanges de données et sur leurs usages des bases de données scientifiques. Deuxièmement, nous avons procédé à l'analyse d'un corpus restreint de documents d'*ArcticNet*, tels que des procès-verbaux, des communications internes, les politiques sur les données de recherche et les rapports annuels. Troisièmement, nous avons pratiqué l'observation participante lors du colloque annuel de 2009, où nous avons participé à des réunions et des ateliers pour nous imprégner du contexte d'*ArcticNet* d'une part, et tenter de gagner la confiance de ses acteurs d'autre part.

Notre corpus d'analyse s'est centré sur les onze entrevues menées auprès des chercheurs d'*ArcticNet*. Les documents d'*ArcticNet* que nous avons lus n'ont pas été inclus dans notre corpus d'analyse. Plutôt, ils ont été considérés comme des éléments contextuels à mobiliser au besoin dans la discussion des résultats. De même, l'observation participante dans les activités d'*ArcticNet* a aussi été considérée comme une source d'information secondaire, notamment parce qu'une minorité de ces activités a été enregistrée. Surtout, dans ces activités, nous n'avons jamais explicitement annoncé les buts de notre recherche et de notre méthode aux participants. Ainsi, pour des raisons à la fois pratiques et éthiques, l'observation participante a principalement servi à nous imprégner du monde d'*ArcticNet* et à nous fournir des éléments contextuels pour l'analyse.

#### 3.4.2 Les entrevues semi-dirigées

Tel que mentionné, l'essentiel des données collectées pour l'analyse provenait d'entrevues réalisés avec des acteurs du réseau *ArcticNet*. Notre stratégie d'échantillonnage a consisté à interviewer des gens travaillant dans la recherche et prenant part d'une manière ou d'une autre à des échanges de données. Il s'agissait donc d'un échantillonnage de type non probabiliste, c'est-à-dire qu'il a été constitué « en fonction des caractéristiques précises que le chercheur veut étudier » (Deslauriers et Kérisit, 1997: 97).

Notre échantillon répondait à deux critères. Premièrement, il devait comporter des interviewés des trois domaines scientifiques d'*ArcticNet* (sciences naturelles, sociales et de la santé) et deuxièmement, représenter les différentes fonctions professionnelles présentes dans *ArcticNet* (professeurs, étudiants, professionnels de recherche). Notre stratégie de constitution de

l'échantillon était à la fois « intentionnelle » (nous avons choisi et contacté directement les sujets) et par « boule de neige » (les sujets rencontrés nous ont recommandé à d'autres sujets potentiels) (Deslauriers et Kérisit, 1997; Mayer et Ouellet, 1991: 386).

Comme nous souhaitions vivement découvrir les contextes scientifiques et organisationnels dans lesquels évoluaient les chercheurs de l'Arctique et que nous désirions laisser la place à la subjectivité et à l'expression de nos informateurs, nous avons opté pour un design d'entrevue de type semi-dirigé. Rappelons-le, les entrevues semi-dirigées sont habituellement organisées autour de quelques questions ouvertes destinées à susciter la réflexion et l'expression des interviewés. L'intervieweur n'hésite pas à demander des éclaircissements sur certaines réponses et à poser des questions impromptues pour relancer l'échange.

Cet outil de collecte présentait plusieurs qualités. D'abord, il était relativement facile et rapide à mettre en place, ce qui était essentiel compte tenu des délais d'une recherche de maîtrise. Ensuite, le recours aux entrevues semi-dirigées nous permettait une grande souplesse dans la conduite de notre terrain de recherche. Si l'interviewé mentionnait une piste intéressante, notre méthode nous permettait de poser de nouvelles questions pour explorer ce filon.

L'entrevue semi-dirigée comportait aussi sa part d'inconvénients. Au-delà du temps supplémentaire à consacrer à la retranscription, le véritable désavantage des entrevues semi-dirigées nous apparaissait être l'imprévisibilité des résultats. Les interviewés étant laissés relativement libres de répondre et d'aborder les thèmes de leur choix, il était un risque de se retrouver avec un corpus plus diffus que si nous avions eu recours à un questionnaire fermé.

Le recours aux entrevues semi-dirigées comme principal instrument de collecte posait aussi quelques interrogations d'ordre épistémologique. Comme le souligne Poupart (1997), l'entrevue est une voie royale pour comprendre les représentations, les pratiques et les enjeux que vivent les acteurs. Aussi riche soit-elle, Poupart indique aussi que cette méthode ouvre la porte aux opinions et aux jugements peu fondés. Le danger est grand pour les chercheurs de confondre faits, convictions, impressions et réalités. S'il s'agissait d'un risque réel, nous considérons que nos ancrages ontologiques et épistémologiques étaient tout à fait appropriés pour intervenir dans un contexte où la subjectivité des acteurs était mise de l'avant. Cependant, le choix d'une épistémologie et d'une méthodologie subjectiviste exigeait que nous discussions de nos résultats avec nuance et circonspection.

En gardant ces considérations en tête, nous avons entrepris de bâtir un guide d'entretien relativement court, construit autour de trois thèmes centraux : 1) usages et pratiques en matière de partage de données, 2) réflexions sur ces pratiques et 3) place des bases de données scientifiques. S'y sont ensuite ajoutée une introduction, des questions sur le parcours professionnel de l'interviewé et une conclusion. Enfin, nous avons produit une douzaine de questions optionnelles pour relancer la discussion. Après plusieurs versions et un important travail de synthèse, le guide d'entretien tenait sur une seule page. Le guide d'entretien et le formulaire de consentement ont fait l'objet d'une traduction en anglais, dans la mesure où plusieurs informateurs étaient anglophones.

### 3.4.3 Le déroulement de la collecte de données

#### 3.4.3.1 La première ronde d'entrevues

Le recrutement de nos informateurs a été grandement facilité par la collaboration de la gestionnaire de données d'*ArcticNet*, madame Josée Michaud. À l'aide d'une liste de professeurs, d'étudiants des cycles supérieurs et de fonctionnaires qui ont soumis des métadonnées au *Polar Data Catalogue* (PDC), nous avons identifié une série de collaborateurs potentiels pour nos entretiens. Comme plusieurs des chercheurs de la liste étaient basés à Québec, nous avons organisé un court séjour les 21 et 22 octobre pour interviewer ces personnes.

Durant notre séjour à Québec, nous avons réalisé cinq entrevues d'une durée moyenne de trente minutes. Malheureusement, l'horaire très chargé de nos répondants nous a contraint à réduire le temps d'entrevue à trente minutes, une condition *sine qua non* pour recruter nos informateurs. À l'évidence, des entrevues plus longues auraient été plus riches. Cependant, le contenu des interviews s'est révélé suffisamment fécond pour l'analyse.

Cette première ronde d'entrevue nous a permis donc de rencontrer un professeur, trois étudiants et une professionnelle de recherche de l'Université Laval et de l'Institut national de recherche scientifique – Centre Eau, Terre et Environnement, tous œuvrant dans les sciences naturelles.

#### 3.4.3.2 La seconde ronde d'entrevues

Suite à notre premier contact avec le terrain, nous avons procédé à quelques changements sur notre guide d'entretien. Les questions qui n'avaient pas su susciter l'expression et la réflexion

ont été retirées ou remplacées par de nouvelles. Nous avons procédé aussi à la traduction en anglais du formulaire de consentement et du guide d'entretien pour les interviews avec les répondants anglophones.

Après quelques tentatives infructueuses de mener une deuxième ronde d'entretien à Québec, nous avons décidé de concentrer nos efforts sur la rencontre scientifique annuelle d'*ArcticNet* (ASM2009), qui a eu lieu du 8 au 11 décembre 2009 à Victoria, en Colombie-Britannique. Ce colloque constituait pour nous une double opportunité. Premièrement, nous avons été invités à y présenter une affiche sur nos recherches, ce qui était une excellente manière de faire connaître nos travaux en dehors de notre cercle restreint de nos collègues. Deuxièmement, la participation à ce colloque nous offrait la chance inestimable de contacter et d'interviewer des chercheurs et des professionnels des provinces anglophones. Également, le fait d'être présent pendant quelques jours sur le terrain nous a permis de vivre et de comprendre les enjeux de la recherche nordique et de ses organisations. Nous y avons mené plusieurs séances d'observation participante durant les conférences, les ateliers ou les événements sociaux.

En faisant connaissance avec quelques-uns des quatre cents congressistes, nous sommes parvenus à recruter six interviewés, deux professeurs, trois étudiants des cycles supérieurs et un professionnel de recherche. Parmi ceux-ci, trois répondants venaient des sciences naturelles, deux des sciences sociales, et une adoptait une approche multidisciplinaire (sciences naturelles et sociales). Quatre entrevues étaient en anglais, deux en français. Encore ici, les entrevues ont duré trente minutes en moyenne, en raison de la faible disponibilité des congressistes.

Ce second contact avec le terrain de recherche a été extrêmement enrichissant pour de nombreuses raisons. Entre autres, les entretiens avec les chercheurs des sciences sociales nous ont donné une perspective très différente sur les enjeux du partage de données et ont grandement stimulé notre réflexion. Aussi, en nous immergeant dans le contexte scientifique et organisationnel d'*ArcticNet*, nous avons développé une connaissance expérientielle de ce milieu, laquelle s'est avérée très importante pour guider notre processus d'analyse des résultats.

#### 3.4.3.3 Le sommaire de la collecte de donnée

Au terme des deux rondes d'entretiens, nous avons réalisé onze entrevues, d'une durée moyenne de trente minutes. Parmi les interviewés, huit provenaient des sciences naturelles, deux des sciences sociales et une adoptait une approche multidisciplinaire. L'absence de chercheurs du

domaine de la santé dans notre échantillon rendait notre portrait moins complet. Par contre, ce constat a été nuancé par la grande variété des disciplines auxquelles les chercheurs rencontrés s'associaient : océanographie, physique, biologie, écologie, études autochtones, anthropologie et géographie. Cette grande diversité d'ancrages disciplinaires et la pluralité des méthodes de recherche nous ont assuré davantage d'occasions de réflexions et de questionnements lors de l'analyse et de la discussion de nos résultats. En outre, nous avons réussi à interviewer des gens occupants différents postes dans leurs équipes de recherche : trois professeurs, six étudiants et deux professionnels de recherche. Cette diversité des profils nous a permis de mettre en relief les propos des interviewés et d'obtenir des points de vue différenciés selon la situation professionnelle de nos informateurs.

### 3.5 L'analyse des données

#### 3.5.1 L'approche générale

Notre procédure d'analyse s'inspire grandement de l'approche mixte de Miles et Huberman (1984). Cette approche propose une méthode générale inspirée de la théorie ancrée (Glaser et Strauss, 1967) et un éventail d'outils pour les démarches de recherche qualitative. Une des idées majeures défendues par Miles et Huberman est que l'analyse des données est indissociable de la cueillette de celles-ci. Plutôt que de se succéder chronologiquement, la cueillette de données et l'analyse sont des processus concomitants qui s'enrichissent mutuellement.

À la différence d'une démarche de théorisation ancrée pure, l'approche mixte de Miles et Huberman permet au chercheur de commencer son analyse avec un cadre conceptuel préexistant (Miles et Huberman, 1984: 28-33). Dans notre cas, nous recourons à la notion des régimes d'échange et à leur typologie (Milanovic, Pontille et Cambon-Thomsen, 2007), qui a fait office de « grille de codage initiale » (Miles et Huberman, 1984: 57).

Tant la théorie ancrée que l'approche mixte demandent que les catégories et les concepts émergents soient confirmés ou explicités à la faveur de nouvelles données empiriques (Glaser et Strauss, 1967; Laperrière, 1997; Miles et Huberman, 1984). Un tel processus comportait des limites importantes, notamment en terme d'accessibilité du terrain et de la quantité de données qu'il a été possible de recueillir et d'analyser convenablement dans le cadre d'un mémoire de maîtrise. Compte tenu de ces limitations, notre démarche méthodologique a principalement été guidée par l'agenda d'*ArcticNet* et la disponibilité des interviewés. Ainsi, nous n'avons pas pu



entamer l'analyse alors que nous procédions à la collecte des données, car les contraintes temporelles nous en empêchaient. Nous nous sommes donc limités à l'ajustement de notre guide d'entretien entre les deux rondes d'entrevue, pour mieux cibler notre objet de recherche. Notre démarche a alors été plus linéaire, l'analyse ne commençant qu'au moment où la collecte des données a été terminée.

En suivant Miles et Huberman (1984), toutes les entrevues ont été transcrites intégralement. Ensuite, une première lecture a été effectuée en usant des codages issus du cadre conceptuel initial (dans notre cas, les régimes d'échange). Pour les situations où les concepts du cadre théorique s'avéraient incomplets ou ambigus, de nouveaux codes ont été créés. Au terme du premier codage, un codage de deuxième degré (*pattern coding*) a été effectué à partir des régularités et des thèmes qui ont émergé lors du codage initial (Miles et Huberman, 1984: 67-69).

À ce stade, Miles et Huberman suggèrent de recourir à des matrices, des grilles, des tableaux, des réseaux conceptuels ou des graphiques pour articuler les thèmes émergents les uns aux autres. En accord avec le principe que l'analyse et l'enquête de terrain sont indissociables, ces schémas nouvellement formés sont ensuite détaillés au profit d'une autre collecte de données. Au terme de cette collecte, une dernière ronde de codage peut être effectuée et les schémas sont finalisés.

Comme nous l'avons précisé plus haut, il nous a été impossible d'effectuer collecte et analyse de manière simultanée. Ainsi, nous n'avons pas eu la chance de pouvoir valider et approfondir les thématiques émergentes lors d'une autre collecte de données. Cependant, nous avons recouru à l'approche mixte de Miles et Huberman (1984) pour leurs nombreuses méthodes proposées pour la catégorisation des données de recherche. Entre autres choses, la méthode de Miles et Huberman se distinguait par l'emploi de moyens graphiques (cartes mentales, tableaux, schémas, métaphores, etc.) pour organiser et analyser les codages en catégories. À notre avis, il s'agissait d'une excellente manière de décloisonner notre démarche et de faire intervenir notre inventivité dans la recherche, tout en restant dans les balises d'un cadre méthodologique relativement commun et accepté.

En janvier et février 2010, nous avons entrepris de retranscrire intégralement les 5 heures d'enregistrement d'entretiens. De ces enregistrements, 120 pages de *verbatim* ont été produites<sup>8</sup>.

### 3.5.2 Le déroulement de l'analyse

Nous exposons notre démarche d'analyse en la divisant en deux étapes successives : le codage des *verbatim* et la catégorisation.

#### 3.5.2.1 Le premier codage

Le codage des *verbatim* a été réalisé à l'aide du logiciel d'aide à l'analyse qualitative *NVivo 7*. Si *NVivo* offrait un large éventail de fonctions d'analyse, notre usage du logiciel s'est limité au codage et à l'organisation des codes.

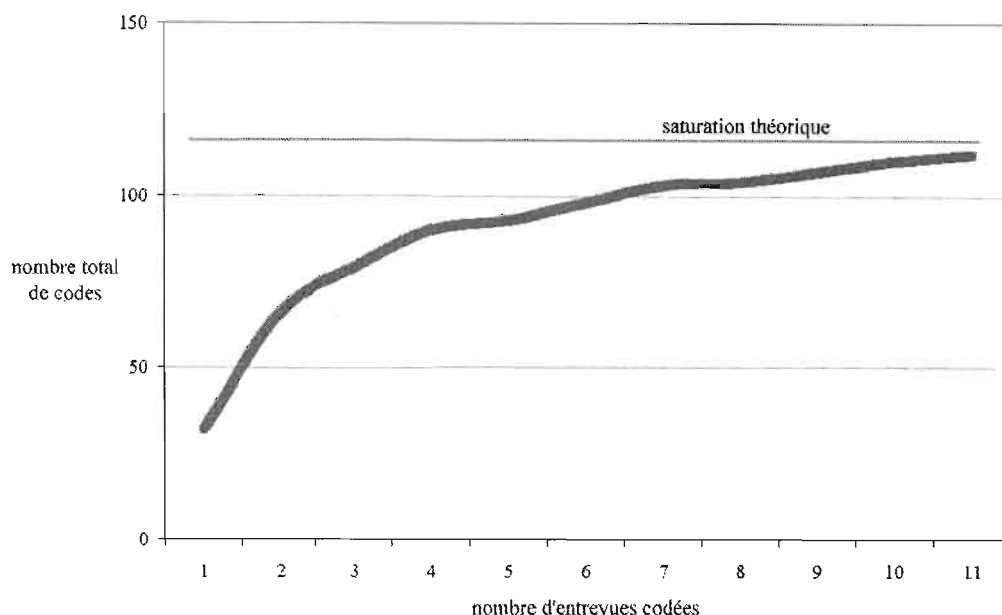
Durant ce premier codage, notre objectif était principalement la *description* de notre corpus. Concrètement, cela impliquait de lire attentivement le texte et de se poser la question « qu'est-ce qu'il y a ici? » à tout moment (Paillé, 1994: 154). La réponse à cette question formait un nouveau code, à moins qu'un code très similaire existe déjà. Dans tous les cas, l'extrait de texte en question était associé au code.

De cette manière, le codage des trois premiers entretiens a généré environ trente nouveaux codes pour chaque entrevue. Par contre, la création de nouveaux codes a considérablement ralenti pour les entretiens suivants, car la plupart des sujets abordés par les interviewés pouvaient être rattachés à des codes déjà existants. Le rythme de création de nouveaux codes a connu une nouvelle baisse dans le codage des derniers entretiens, ce qui peut être considéré comme un indice de l'atteinte d'une saturation théorique<sup>9</sup>. La figure 3.1 illustre l'atteinte de la saturation théorique autour du codage de la septième entrevue.

---

<sup>8</sup> Nous avons eu la chance d'être aidés dans cette tâche par madame Lucie Enel, assistante de recherche de notre directrice madame Millerand. Madame Enel s'est chargée de la retranscription des sept entrevues en français alors que nous avons retranscrit celles en anglais.

<sup>9</sup> « La saturation signifie qu'aucune donnée additionnelle ne serait utile au sociologue pour développer davantage une catégorie » (Glaser et Strauss, 1967: 61).



**Figure 3.1** Nombre total de codes et nombre d'entrevues codées

Après le codage des premiers *verbatim*, un important travail d'organisation des codes a été entrepris. Premièrement, les codes ont été rassemblés en groupes tels que « facilitateurs du partage », « inhibiteurs du partage », « règles et systèmes de fonctionnement », etc. Ce travail d'organisation a été ensuite répété après le codage de chaque *verbatim*. Deuxièmement, pour faciliter leur repérage et leur manipulation, un code unique de trois lettres a été attribué à chaque code.

Ainsi, au terme de cette première ronde de codage, nous nous sommes retrouvés avec cent neuf codes représentant différentes facettes du partage de données scientifiques. Au total, les codes ont été associés à sept cent trente-cinq citations dans les *verbatim* d'entrevue. La plupart des codes ne référaient qu'à un ou deux passages des retranscriptions. Par contre, certains codes pointaient vers plus d'une dizaine de citations dans notre corpus.

### 3.5.2.2 La catégorisation et la schématisation

Après ce premier travail de codage, nous avons entrepris d'organiser, de structurer et de schématiser les codes que nous avons générés. Il s'agissait d'une étape importante, dans laquelle

nous nous sommes efforcés de passer de la simple description de la situation (le codage initial) à une conceptualisation plus raffinée (tableau, typologie, schéma, etc.).

À cette étape, la principale problématique à laquelle nous nous sommes heurtés était la grande quantité de codes et de références à traiter et à analyser. Comment pouvions-nous analyser convenablement cette masse d'information?

Notre solution à cette problématique a été de recourir à des règles de classement et d'organisation qui nous ont permis de mettre en évidence les codes qui se rapprochaient le plus de nos interrogations tout en reléguant les informations périphériques à l'arrière-scène. Évidemment, une telle démarche de hiérarchisation des codes comportait son lot d'arbitraire et d'occasions manquées. Cependant, il s'agissait d'un pari intéressant, compte tenu de la plus grande concision et de la meilleure intelligibilité que nous permettait cette organisation des codes.

En procédant à la lecture répétée la liste des codes et de leurs citations associées, nous avons regroupé les codes autour des trois régimes d'échange décrits par Milanovic, Pontille et Cambon-Thomsen (2007). Certains codes ont été associés à plus d'un régime et plusieurs codes n'ont été rattachés à aucun régime.

Les codes n'ayant pas été associés à un régime ont été ensuite examinés et évalués en vertu des deux questions suivantes : « ce code rend-il compte d'un élément contextuel majeur dans le partage de données? » et « ce code fait-il état de pratiques ou de réflexions sur le partage de données qui ne peuvent pas être comprises dans le cadre de la théorie des régimes d'échange? ». Si la réponse aux deux questions était non, nous retirions le code de notre liste, nous réservant le droit de nous y référer plus tard au besoin. Dans les cas où il s'agissait d'un code se référant à un élément de contexte, nous l'avons ajouté dans le groupe « contexte ». Enfin, dans les cas où nous étions en présence d'une situation qui ne pouvait être expliquée par la théorie des régimes d'échange, nous avons regroupé ces codes dans le groupe « au-delà des régimes ».

Au terme de cette étape de catégorisation, nous sommes arrivés à représenter graphiquement et métaphoriquement l'organisation des codes. Cette étape a été marquée par la créativité, alors que nous avons esquissé des cartes mentales et des analogies susceptibles d'illustrer la topographie conceptuelle du partage de données. Une des analogies que nous avons retenues a été celle de la mappemonde, dans laquelle on retrouvait cinq grands *continents*-

*groupes* : le régime de coopération, le régime de don, le régime de sous-traitance, le contexte et « au-delà des régimes ». En poursuivant cette analogie géographique, dans chaque continent se trouvaient des *pays codes*, dont certains se situaient aux frontières de deux ou plusieurs continents. Ces exercices d'organisation et de visualisation nous ont permis de distinguer les données centrales à notre projet de celles qui faisaient figure d'éléments contextuels, et d'écarter les données insignifiantes pour notre recherche.

L'étape suivante a été de s'interroger plus en profondeur sur chacun des trois groupes de codes associés aux régimes d'échange : « comment les reconnaît-on? », « quels sont leurs caractéristiques? », « comment ces régimes prennent-ils sens pour leurs acteurs? », « qu'est-ce qui distingue ce régime des autres? » ou encore « quels liens peut-on tracer entre chacun des régimes? ». Pour ce faire, nous sommes retournés directement lire les extraits tirés du corpus et sommes allés chercher des informations complémentaires dans les codes conceptuels.

Les codes situés dans le continent « au-delà des régimes » ont été analysés distinctement et avec un grand intérêt, parce qu'il s'agissait selon nous des meilleures pistes pour développer et dépasser notre cadre théorique initial. Pour l'analyse de ces codes particuliers, nous avons procédé de la même façon qu'avec les *continents-régimes*, soit en retournant à l'examen du corpus. Nous sommes donc parvenus à cerner les caractéristiques et les particularités de ces codes, en recourant aux éléments contextuels au besoin.

À ce stade, l'étape d'analyse et de théorisation arrivait à sa fin. La dernière phase importante a consisté à mettre en relations nos résultats d'analyse avec les éléments conceptuels retenus dans notre cadre théorique (nouvelle sociologie des sciences et des technologies, sociologie des usages, informatique sociale).



## CHAPITRE IV

### RÉSULTATS

Tel que mentionné dans le chapitre précédent, notre démarche d'analyse a été marquée par une grande quantité de données. Les résultats que nous présentons dans les prochaines pages ne représentent donc qu'une fraction de la masse de données collectées. Il s'agit des résultats les plus pertinents compte tenu de notre problématique et de nos orientations théoriques.

Le présent chapitre est divisé en cinq sections. Premièrement, nous présentons le contexte général dans lequel s'effectue le partage des données. Deuxièmement, nous consacrons trois sections à la description des pratiques de partage de données. Chacune de ces sections est organisée autour d'un régime d'échange (régimes de coopération, de don et de sous-traitance; Milanovic, Pontille et Cambon-Thomsen, 2007). Finalement, la dernière section est consacrée aux cas d'exception et aux pistes d'analyse qui en découlent.

#### 4.1 Le contexte général du partage de données

Si nos interviewés ont comme point commun d'œuvrer dans le domaine de la recherche en Arctique, leurs méthodes, terrains, données et ancrages disciplinaires sont très différents. Il n'y a donc pas un seul, mais plusieurs contextes dans lesquels s'effectue le partage de données, selon les disciplines et les communautés de recherche. Afin de brosser un portrait rapide de la situation, nous mettons en évidence quatre éléments contextuels majeurs : la recherche sur l'Arctique, la nature des données collectées, la multidisciplinarité et la perception du caractère inévitable du partage des données.

#### 4.1.1 Le contexte de la recherche sur l'Arctique

Depuis quelques années, les questions liées à l'Arctique occupent une place grandissante dans la sphère publique. Les enjeux liés à l'exploitation des ressources naturelles, aux disputes territoriales entre les nations, à la militarisation du Nord et aux changements climatiques font de l'Arctique un sujet chaud dans les médias canadiens.

Au même moment où l'Arctique devient un sujet d'actualité, la recherche scientifique dans le Nord canadien est en pleine expansion :

*Maintenant, c'est la mode, c'est comme la dernière frontière [en parlant de la recherche arctique]. Je me souviens, au début, il n'y avait jamais de brise-glace de recherche qui allaient là. L'été passé ou l'été d'avant, il y en avait quatre [...] dans la même région. (Patrice, professeur en physique) <sup>10</sup>.*

La recherche arctique pose des enjeux logistiques de taille. Il s'agit d'un immense territoire très peu peuplé. L'isolement et les conditions météorologiques extrêmes rendent l'établissement de bases de recherche permanentes très difficile. De même, le transport du personnel et du matériel dans les territoires nordiques est extrêmement coûteux. Ce contexte particulier a une influence importante sur le partage de données. Contraints par des ressources limitées et des conditions de collecte difficiles, le partage de données s'impose comme solution pour les chercheurs en Arctique.

*Olivier : Qu'est-ce qui te motive à investir les efforts nécessaires [au partage de données]? Elena : C'est la réalité du travail sur un navire en Arctique. On est très limité en espace, on est très limité en temps aussi. [...] On peut réussir à avoir des vues beaucoup plus globales de ce qui se passe. En réussissant à avoir des données qu'on n'aurait tout simplement pas eu le temps de prendre, ou pas eu l'argent, ou pas eu le personnel nécessaire pour prendre... c'est l'avantage, ça nous permet d'avoir un plus grand nombre de données. (Elena, étudiante en biologie).*

---

<sup>10</sup> Les citations d'extraits de *verbatim* sont identifiées par le pseudonyme de la personne interviewée, son statut (ex : « étudiant » ou « professeur ») ainsi que sa discipline ou domaine d'appartenance (ex. « physique » ou « biologie »).



#### 4.1.2 Les différences dans la nature même des données

Au cours des entrevues, les interviewés nous expliquent que la nature même des données influence les modalités de leur partage éventuel. D'un côté, les données physiques et chimiques de base (ex : température, salinité, nutriments...), qui sont nécessaires aux travaux de nombreux chercheurs, sont partagées d'emblée. L'importance et la centralité de ce genre de données a d'ailleurs conduit *ArcticNet* à mettre en place des politiques visant à en faciliter le partage :

*Au sein d'ArcticNet, dans les designs de la recherche [...] il y a ce qu'on appelle les Core Data. [...] il suffit de faire une demande pour les obtenir.* (Édouard, étudiant en biologie).

De l'autre, les données de sciences sociales, notamment qualitatives, sont plus difficiles à publier et à partager pour des raisons éthiques et de confidentialité :

*In terms of raw data, or research produces, it's generally interview data, which is very difficult to share. [...] A big part of it is confidentiality, that when we go in to do interviews, we fill out the confidentiality forms and it's meant to be solely for that one researcher's use to analyze<sup>11</sup>.* (Thomas, professionnel de recherche en sciences humaines).

Un autre facteur contextuel important du partage de données est celui de l'abondance des données de recherche. Malgré les coûts élevés et les difficultés logistiques liés à la recherche nordique, plusieurs chercheurs nous indiquent qu'ils font face à une abondance de données (un phénomène que plusieurs auteurs ont appelé le déluge de données; Anonyme, 2010; Borgman, Wallis et Enyedy, 2007). Dans ce contexte, on nous indique que le partage est rendu plus facile :

*[...] normalement [...] [l'accès aux données] était réservé aux gens du projet tant que le projet était pas fini [...] Dans les faits, il y en a tellement que même moi, j'arrive pas à toutes les analyser. Donc les données CTD, on les annonce pas à l'international, mais il suffit que tu nous demandes [et] on les donne, les données de physique. »* (Patrice, professeur en physique).

---

<sup>11</sup> Nous avons choisi de citer les extraits de *verbatim* dans leur langue originale, et donc de ne pas les traduire, pour demeurer fidèle au corpus et par souci d'authenticité.

Enfin, en assistant aux présentations et aux ateliers de la rencontre annuelle d'*ArcticNet*, nous sommes sensibilisés à une problématique jusqu'alors complètement absente de notre démarche de recherche. D'une manière ou d'une autre, tous les chercheurs qui travaillent sur l'Arctique doivent interagir avec les populations inuites du Nord canadien. Dans le cas des sciences sociales et des sciences de la santé, ce sont souvent les communautés inuites elles-mêmes qui constituent les objets de recherche. Au moment où le Nord canadien connaît des changements climatiques et sociaux majeurs, plusieurs chercheurs d'*ArcticNet* s'intéressent en effet aux savoirs traditionnels entretenus et développés par les communautés inuites (techniques de chasse et de pêche, connaissance des lieux, alimentation, médecine, etc.).

Les chercheurs des sciences sociales que nous avons interviewés nous révèlent que les savoirs traditionnels inuits constituent des informations très confidentielles et qu'à ce titre, ils doivent traiter ces données avec beaucoup de précaution et de respect :

— *[...] with the char data [un poisson de genre Salvelinus], it's data that's coming from someone's career efforts. It's their jobs, it's what they're paid for, so maybe, if they won't use it, [it] doesn't help their career, they are more willing to part with it. [...] Whereas for traditional knowledge, you're talking of someone's life. Like what they've lived, or what their parents lived, their brother, their friends or their personal experience. It's not [like] you're just walk in and ask for it. [...] You have to really make sure people [are] understanding the project. You have to make sure that they know that they will be respected. That their knowledge will be respected and it won't be shared without their permission.* (Emma, étudiante en biologie et en sciences humaines).

Dans ce contexte, les données collectées dans les communautés inuites sont considérées comme demeurant la propriété des communautés, et celles-ci qui doivent être consultées dans tout partage éventuel des données :

*I've had a lot of people asking me for the traditional knowledge from King's Bay [lieu fictif] outside the community, but it's not mine to share. That's the community's knowledge. So even though I've been given permission to document it and use it, I can't just offer it.* (Emma, étudiante en biologie et en sciences humaines).

#### 4.1.3 La collaboration multidisciplinaire

L'interdépendance entre les chercheurs dans les équipes multidisciplinaires constitue un autre élément contextuel majeur que nous relevons dans nos entrevues. En particulier en sciences

naturelles, on reconnaît aisément que les données des collègues sont essentielles à la réussite des projets de recherche de chacun.

*Puisqu'on fait de l'océanographie, c'est à peu près impossible de faire de l'océanographie dans un autre contexte parce que [...] les biologistes ont besoin des données chimiques, de nutriments, etc. Les gens qui travaillent sur les nutriments et les biologistes ont tous besoin de données de physique puis si on veut comprendre ce qui se passe dans l'eau, au niveau des organismes etc. [il] faut comprendre les courants [il] faut comprendre... Donc à la base, c'est quelque chose de vraiment très, très multidisciplinaire, donc je vous dirais que c'est un peu nécessaire et naturel d'avoir ce transfert de données là. (Paul, professeur en biologie).*

Cela dit, cette dépendance entre chercheurs n'est peut-être pas aussi réciproque qu'elle ne le paraît. Si l'on reconnaît l'importance centrale des données physiques dans la recherche en biologie, en chimie ou en écologie, les physiciens semblent en revanche avoir peu besoin de leurs collègues d'autres disciplines. Cette citation illustre bien cette disproportion des besoins :

*C'est parce que nous, en physique, on est différent de tous les autres parce que... c'est sûr que je peux utiliser les données des autres, mais les autres ont absolument besoin des données de physique [...]. (Patrice, professeur en physique).*

#### 4.1.4 Une perception commune : le développement du partage de données est inéluctable

Si les pratiques de partage de données et les contextes scientifiques dans lesquels elles s'insèrent varient d'une situation à l'autre, nos interviewés sont unanimes dans leur manière de parler de l'évolution du partage de données et de la collaboration scientifique. Le fait que la science devient de plus en plus collaborative et que le partage de données est une pratique de plus en plus répandue est maintes fois exprimé :

*I think that there will definitely be more [data sharing], I think it's [...] definitely becoming more open and people are [...] not as reluctant and understand the importance of data sharing and having the data open. (Elaine, étudiante en biologie).*

Autrement dit, les chercheurs sont conscients du fait que le partage des données scientifiques n'est plus un scénario d'avenir parmi d'autres, mais qu'il s'agit d'une tendance inéluctable et cela, même s'ils reconnaissent que les critères déontologiques des sciences sociales en particulier posent de sérieuses limites :

*In terms of social sciences, I'm not sure if it can substantially improve because of the limitation of not being able to share on.*  
(Thomas, professionnel de recherche en sciences humaines).

#### 4.2 Le régime de coopération

Tel que mentionné dans notre cadre théorique, nous reprenons la théorie des régimes d'échange développée par Milanovic, Pontille et Cambon-Thomsen (2007) qui propose une typologie en trois régimes (de coopération, de don et de sous-traitance) et qui définit un régime d'échange comme suit : « [Un régime d'échange] qualifie les liens entre les parties prenantes et définit les formes (matérielles, relationnelles et juridiques) de l'échange de manière spécifique » (Milanovic, Pontille et Cambon-Thomsen, 2007: 19).

Le premier élément de la typologie, le *régime de coopération*, est un régime d'échange très peu formalisé qui repose sur la proximité des acteurs et sur un très haut niveau de confiance mutuelle. Dans *ArcticNet*, ce régime semble être principalement à l'œuvre chez les chercheurs en sciences naturelles, lesquels utilisent régulièrement plusieurs jeux de données dans leurs recherches. Le régime de coopération s'actualise également chez les chercheurs de sciences humaines, mais dans une moins grande mesure, leurs données étant protégées par des normes éthiques sévères. Les résultats montrent que ce mode de collaboration est présent tant chez les professeurs que chez les étudiants ou les professionnels de recherche.

Le régime de coopération se distingue par quatre caractéristiques. Premièrement, la proximité entre les partenaires de l'échange est une condition *sine qua non* du régime de coopération. Les partenaires se connaissent bien et ont déjà collaboré par le passé. Deuxièmement, la confiance entre les acteurs occupe un rôle central dans ce régime. C'est par le jeu de confiance mutuelle qu'un partage satisfaisant peut être établi. Troisièmement, le régime de coopération est marqué par la facilité des échanges. Comme les partenaires bâtissent une relation de confiance et de collaboration rapprochée, le partage de données se fait de manière aisée et non problématique. Quatrièmement, la réciprocité des échanges est appréciée, mais elle n'est pas obligatoire dans le régime de coopération. Dans les prochaines lignes, nous examinons chacune de ces caractéristiques.

#### 4.2.1 La proximité des acteurs

Pour nos interviewés, le partage de données s'inscrit avant tout dans des relations de proximité. Par relations de proximité, nous entendons des situations où les acteurs entretiennent ou ont entretenu des relations professionnelles soutenues. En théorie des réseaux sociaux, nous pourrions qualifier ces relations de *liens forts* (Granovetter, 1973). Il peut s'agir de collègues travaillant dans le même département, de collaborateurs de longue date, de chercheurs rencontrés lors des terrains de recherche, etc. Cette citation illustre bien ce sentiment de proximité et d'attachement à une équipe :

*I think... being part of that group... we actually do work fairly closely with a lot of the group because we're at sea together, we're working on a project together. In some cases, I've analyzed samples for other groups as well, so it's data sharing, but we're collaborating with them.* (Elaine, étudiante en biologie).

À la lumière des entrevues, il apparaît que proximité et partage de données vont de pair, comme l'illustre cette citation, qui parle du partage de données comme d'une culture qui se développe dans le temps :

*[...] dans [...] la communauté dans laquelle j'évolue, oui, je crois que ça l'est [la perception de l'importance du partage de données]. Je pourrais pas parler pour d'autres disciplines, je sais pas si c'est aussi bien huilé dans d'autres disciplines que dans la nôtre. On a un long historique de projets, de collaborations rapprochées. Beaucoup d'entre nous, les chercheurs senior, ça fait peut être quinze, vingt ans qu'on est dans des projets comme ça, donc la culture pour nous est là.* (Paul, professeur en biologie).

L'importance de la proximité dans le régime de coopération fait en sorte que l'on gagne beaucoup à agrandir son réseau de recherche. Plusieurs interviewés nous expliquent qu'une grande partie des données proviennent de gens évoluant dans leur réseau professionnel :

*[...] ça fait tellement longtemps que je suis ici que les données que je veux avoir, je sais déjà à qui m'adresser pour les avoir directement. [...] Je sais que [pour] ce que j'ai besoin, je sais déjà qui contacter, qu'est-ce qu'il va avoir comme donnée, où, quand, comment et pourquoi.* (Elena, étudiante en biologie).

#### 4.2.2 L'importance de la confiance

Un des corollaires directs de la grande proximité des acteurs est le rôle central de la confiance pour encadrer les échanges. La confiance est perçue comme une relation qui se bâtit avec le temps et qui facilite le partage entre les chercheurs :

*[...] Comme je disais tout à l'heure, c'est facile de partager des données avec des gens avec qui on a bâti une confiance. C'est peut être ça qui va limiter un peu l'étendue du réseau ou la vitesse à laquelle le partage va se faire : ce jeu de confiance qui doit se créer entre les gens [...].* (Pierre, professeur en écologie).

Une autre interviewée souligne qu'il est important d'avoir confiance dans la justice et la réciprocité des échanges :

*La confiance... [Il] faut avoir confiance avec les gens avec qui on travaille. Savoir que de un, ces personnes vont ramasser de bonnes données, puis [...] qu'un coup que les données vont être ramassées, qu'ils vont être prêts à faire un partage équivalent, donc si eux me font des demandes, mais lorsque moi j'arrive à leur en faire, qu'ils soient prêts à répondre à ce qu'on leur donne.* (Elena, étudiante en biologie).

Par ailleurs, dans le cadre de sa recherche sur les communautés inuites, une étudiante nous fait part du fait qu'elle est retournée dans les communautés inuites étudiées pour leur remettre les retranscriptions d'entrevues ainsi qu'un résumé de sa recherche. Pour elle, ce geste a surtout servi à cimenter la relation de confiance qu'elle avait établie avec ces communautés :

*So, just as far as maintaining the relationship and trust, it was more about that than [...] that they want that piece of paper.* (Emilie, étudiante en sciences humaines).

#### 4.2.3 La facilité des échanges

La grande proximité entre les chercheurs et les relations de confiance qui se nouent entre eux rendent les échanges de données très aisés. Les interviewés nous indiquent qu'il suffit le plus souvent de contacter la personne responsable par courriel ou par téléphone pour obtenir les données. Celles-ci sont alors partagées au moyen de technologies relativement simples comme le courriel, les CD ou les sites FTP.

*[...] Si tu veux aller voir le chef de mission puis lui demander : « Je pourrais avoir les données sur un CD? » Puis il te les donne ou ensuite, tu peux aller chercher aussi sur le site Internet. Eux, ils les*

*diffusent sur le site internet puis pour ArcticNet, tu as juste à lâcher un coup de fil à Guy Robert [nom fictif] puis il va t'envoyer toutes les affaires pour... les données de sel et de température. (Édith, étudiante en biologie).*

*[...] avec les gens avec qui je travaille depuis longtemps [...] c'est très informel, donc on a beaucoup échangé de données. [...] Quelqu'un envoie un courriel pour demander un jeu de données particulier puis s'il n'y a pas de problème, normalement, on l'envoie. (Paul, professeur en biologie).*

Cette manière d'échanger les données qui repose sur la facilité, la proximité et la confiance semble être de loin la plus répandue chez nos interviewés. Tous parlent de ces échanges en des termes positifs. Il s'agit d'échanges entre pairs, au bénéfice de tous, et dans le cadre de l'appartenance à une équipe, un champ ou une discipline. L'absence d'ambiguïté, de formalisme ou de dispositifs techniques complexes à maîtriser semble rendre le partage encore plus facile.

Dans ce type de régime d'échange, les bases de données ne sont pas ou peu mobilisées, du moins directement. En effet, les chercheurs étant déjà engagés dans des relations de collaboration rapprochées, l'usage de dispositifs techniques comme des bases de données en ligne apparaît superflu, d'autant plus s'il s'avère complexe :

*[...] aller consulter de façon simple et efficace les données des autres, c'est pas toujours facile! (Paul, professeur en Biologie).*

#### 4.2.4 Une réciprocité appréciée, mais facultative

Comme dernière caractéristique notable du régime de coopération, nous relevons la dimension particulière que prend la réciprocité des échanges. En prenant un peu d'avance sur la section suivante, rappelons simplement que la réciprocité est au cœur du régime de don. Par contre, elle n'est pas tout à fait absente dans le régime de coopération.

Si l'on nous révèle que le partage s'effectue facilement, avec bonne volonté et « de façon assez libre » (Paul, professeur en biologie), il demeure qu'un retour de la part des demandeurs de données est généralement apprécié. Par contre, cette pratique n'est pas forcément exigée ou même attendue. Elle tiendrait davantage d'une marque de politesse destinée à remercier le fournisseur des données et viserait par ailleurs à renforcer la relation de coopération :

*[...] les données de CTD, ils les donnent à tout le monde, ils peuvent faire ce qu'ils veulent avec. Bon, s'ils sont gentils, ils vont*

*m'avertir puis ils vont m'offrir des collaborations sur l'interprétation, mais ils pourraient ne pas le faire, il n'y a aucune obligation.* (Patrice, professeur en physique).

En somme, les témoignages des interviewés dans le régime de coopération font état de pratiques s'inscrivant dans la familiarité et la collégialité. Les échanges qui s'y développent s'inscrivent dans des liens de collaboration où la générosité et la bonne entente mutuelle encadrent et simplifient la circulation des données. Cependant, pour prendre place, ce régime de coopération exige que les partenaires de l'échange établissent préalablement une relation de confiance et de reconnaissance (Milanovic, Pontille et Cambon-Thomsen, 2007: 21-23).

#### 4.3 Le régime de don

Le deuxième régime d'échange est le *régime de don*. Il s'agit d'un régime d'échange reposant sur la réciprocité des dons entre acteurs (Milanovic, Pontille et Cambon-Thomsen, 2007: 23). Dans *ArcticNet*, cette forme d'échange est présente principalement dans les sciences naturelles, lesquelles peuvent partager leurs données sans restrictions éthiques. Nos résultats ne font pas état de ce régime dans les sciences sociales, qui doivent protéger la confidentialité de leurs informateurs. Le régime de don implique autant des chercheurs expérimentés que des étudiants et des professionnels de recherche.

En comparaison avec le régime de coopération que nous venons de discuter, le régime de don renvoie à des échanges plus distants et davantage négociés entre les partenaires. Le partage initial d'un jeu de données appelle un contre-don pouvant revêtir différentes formes.

Une première caractéristique du régime de don est qu'il peut faire intervenir des partenaires distants. Sans être complètement étrangers l'un à l'autre, les partenaires du régime de don n'ont pas besoin d'avoir établi une relation préalable comme dans le régime de coopération. Une deuxième caractéristique est celle de la réciprocité des dons. Le *quid pro quo* entre les partenaires peut consister en l'échange d'un jeu de données contre une mention d'auteur dans une publication future, par exemple. Notons également que dans ce régime, les étudiants gradués peuvent être appelés à jouer un rôle d'intermédiaire ou de « monnaie d'échange » dans les partenariats entre chercheurs. Voyons comment les caractéristiques de ce régime se manifestent dans *ArcticNet*.



#### 4.3.1 Des partenaires distants

La première caractéristique importante du régime de don est qu'il fait intervenir des acteurs qui n'ont pas forcément développé de lien au préalable. Nos entretiens révèlent par exemple que les demandes de partage de données peuvent provenir d'autres laboratoires ou d'autres pays. Cette plus grande distance relationnelle entre les partenaires force la prudence :

*On peut recevoir des sollicitations de personnes qu'on connaît pas nécessairement. Et dans ces cas-là, il faut faire attention. (Édouard, étudiant en biologie).*

Contrairement à la simplicité et la facilité des échanges qui ont cours dans le régime de coopération, le fait d'interagir avec des partenaires inconnus amène les chercheurs à demander des garanties ou des renseignements supplémentaires :

*Il n'y a pas vraiment d'entente formelle, de choses signées, mais c'est des échanges de e-mails. Je vais recevoir un e-mail qui demande telle chose puis après ça, si le e-mail qui fait la demande est pas très, très précis, on va juste essayer d'avoir des précisions sur l'utilisation des données puis je vais demander à ce que l'utilisation des données soit pas diffusée ou que si quelqu'un veut utiliser les données à une tierce partie, que cette tierce partie m'en fasse directement la demande, à moi, histoire de savoir un peu ce qui arrive. (Paul, professeur en biologie).*

Cela dit, si le régime de don peut faire intervenir des partenaires externes, il peut aussi qualifier certains échanges qui ont cours entre personnes à l'intérieur d'un groupe ou d'une organisation.

#### 4.3.2 La réciprocité des dons

La caractéristique centrale du régime de don est la réciprocité dans les échanges. Dans leur théorisation, Milanovic, Pontille et Cambon-Thomsen choisissent de recourir à la sociologie du don (Mauss, 1950) pour illustrer les échanges mutuels qui ont lieu dans les biobanques. Pour l'essentiel, cela suppose un jeu relationnel dans lequel une première personne fait un don à une seconde, à la suite de quoi cette dernière se sent obligée et offre un contre-don en retour.

À travers les entrevues, cette logique de réciprocité des dons est évoquée à plusieurs reprises. La forme la plus commune reste sans doute l'échange « données contre publications » :

*I think it is just generally understood that if [you] use someone else[']s] data they should be a co-author or acknowledged extensively in your paper. (Emma, étudiante en biologie et en sciences humaines).*

Souvent, cette forme d'échange implique aussi une invitation à des projets d'articles, ce qui implique de fournir des données et de participer à l'écriture.

Dans d'autres cas, on utilise la mention d'auteur comme bien symbolique dans l'échange, sans que le fournisseur de données ait nécessairement participé au processus de rédaction :

*Il y a certaines personnes avec qui on va partager des données puis très rapidement, il y a un retour sur l'investissement. C'est-à-dire que moi, j'investis pour structurer ma base de données, j'investis pour bien décrire, je réponds rapidement à la demande de la personne et rapidement, si j'ai un retour là-dessus, c'est-à-dire « Oh, on est en train de travailler sur un manuscrit, il y a une publication en cours, ton nom... tu es coauteur sur une présentation orale » puis ensuite « Oh, tu te retrouves coauteur sur un article. » (Pierre, professeur en écologie).*

Une autre forme commune de contre-don est l'inclusion du nom du fournisseur des données dans les remerciements des présentations orales et des articles scientifiques. Cette forme de réciprocité semble convenir aux situations où le rôle des données en question reste relativement marginal.

Comme dans le régime de coopération, le rôle des bases de données est marginal dans le régime de don. En effet, le régime de don implique un lien direct entre des partenaires, sans passer par une tierce ressource comme une base de données. De plus, le jeu de *quid pro quo* assure une récompense à court terme aux scientifiques, ce qui détonne avec la contribution « altruiste » qu'on leur demande de faire dans les bases de données.

*[...] on reconnaît l'avantage de partager, l'avantage de bien archiver, structurer, etc. J'en suis très conscient, mais ça prend du temps. ça prend de l'énergie puis comme chercheur, la pression, c'est de publier à court terme, que tu sois très productif à court terme. Donc on voit très bien les bénéfices à long terme [des bases de données] mais y'a pas un organisme subventionnaire qui me donne une subvention pour 20 ans. (Pierre, professeur en écologie).*

#### 4.3.3 Les étudiants comme intermédiaires

Enfin, un des interviewés nous indique que ces échanges interviennent beaucoup par l'intermédiaire des étudiants des cycles supérieurs. Leur mobilité semble en effet jouer un rôle central dans l'établissement de relations de collaboration entre les chercheurs.

*[...] je vais souvent agir comme cosuperviseur ou comme conseiller puis je vais fournir une partie des données que l'étudiant va utiliser pour finir son projet de thèse ou de mémoire puis ça va souvent mener à des publications avec plusieurs coauteurs. A ce moment-là, en partageant mes données, je deviens coauteur sur un manuscrit. Donc ça, c'est très commun. [...] Jusqu'à maintenant, le partage s'est surtout fait à travers des étudiants gradués [...] qui complètent des projets et qui vont chercher des données dans plus qu'un laboratoire ou dans plus qu'une équipe. (Pierre, professeur en écologie).*

En définitive, le régime de don renvoie à des pratiques d'échange relativement disparates qui ont comme point commun de s'appuyer sur un jeu de dons bilatéraux. Contrairement au régime de coopération, les partenaires de l'échange peuvent être plus distants. Cependant, contrairement au régime de sous-traitance que nous discutons ci-dessous, ces échanges ont encore lieu de manière relativement informelle et peu codifiée.

#### 4.4 Le régime de sous-traitance

Le dernier régime d'échange est le *régime de sous-traitance*, qui peut être défini en ces termes : un régime d'échange basé sur une offre commerciale de services qui est hautement formalisée (contrats, traces écrites, documents juridiques, etc.; Milanovic, Pontille et Cambon-Thomsen, 2007).

Contrairement aux régimes de coopération et de don, le régime de sous-traitance ne semble pas être très commun dans *ArcticNet*. Il faut dire que les chercheurs d'*ArcticNet* sont surtout des professeurs d'université subventionnés par les agences de financement fédérales et provinciales, qui travaillent peu avec le milieu de la recherche privée. En choisissant ce terrain d'enquête, nous réduisons nos chances de rencontrer des chercheurs négociant des contrats commerciaux de prestation de service, en l'occurrence, des contrats de partage ou d'échange de données de recherche. Cependant, certaines pratiques évoquées par les interviewés s'approchent de ce régime de sous-traitance.

Dans *ArcticNet*, ce régime d'échange impliquait autant des chercheurs des sciences sociales que des chercheurs des sciences naturelles. L'intérêt des communautés locales pour les données des sciences sociales les concernant donne lieu à des ententes entre les chercheurs et les communautés. Les chercheurs en sciences naturelles sont aussi appelés à partager avec des gouvernements, des communautés et des compagnies privées; dans le cadre d'études d'impact, par exemple. Enfin, notre échantillon montre que ce régime peut être négocié autant par les étudiants que les professeurs.

Une première caractéristique du régime de sous-traitance est qu'il met en relation des partenaires hétérogènes. Lorsqu'un chercheur négocie avec des fonctionnaires ou des compagnies privées, ce sont des mondes sociaux très différents qui se rencontrent. Les objectifs, les méthodes, le jargon et les valeurs ne sont pas les mêmes. Une deuxième caractéristique du régime de sous-traitance est qu'il implique une grande spécialisation des tâches entre les chercheurs. Selon les dispositions de l'entente, certains chercheurs peuvent se spécialiser dans l'acquisition, l'archivage ou le partage de données. Enfin, dans le cadre de ce régime, les accords écrits et signés prennent le pas sur les ententes de bonne foi qui prévalent dans les deux autres régimes.

#### 4.4.1 Des partenaires hétérogènes

Dans le régime de sous-traitance, l'écologie des partenaires de l'échange se complexifie. Cela contraste radicalement avec les deux premiers régimes, lesquels ne font intervenir que des chercheurs.

Ce régime renvoie principalement aux divers échanges entre les chercheurs et le gouvernement, incluant ses multiples composantes organisationnelles. Au Canada, la recherche en territoire nordique est étroitement encadrée par le gouvernement fédéral. De plus, le design même du réseau de centres d'excellence *ArcticNet* prévoit la participation active de fonctionnaires. Plusieurs interviewés nous font part des échanges et des collaborations qu'ils mènent avec différentes instances gouvernementales, lesquels s'effectuent principalement dans une logique de financement et d'octroi de permis. En échange d'une subvention et d'un permis permettant à la recherche de se dérouler dans une zone protégée par exemple, les chercheurs se doivent, en retour, de partager leurs données.

*It was a condition of my funding [de collecter et de partager des données]. Because my funding comes from the fisheries joint*

*management committee, they represent the concerns of the people and the concerns for this project was that there is something happening to the char [and] they don't know what. [...] So I had to say that I would collect a huge sweep of data in order to help them out and they help me by giving me funding.* (Emma, étudiante en biologie et en sciences humaines).

Dans d'autres cas, l'échange s'effectue avec des partenaires du secteur privé, notamment dans le cadre d'études d'impact.

*ArcticNet a fait une entente avec Billio Canada [compagnie fictive]. Il a signé une entente où cet été, on a pris des données pour eux et ils étaient à bord. C'est parce qu'ils doivent faire des études d'impact pour leur plateforme de forage. L'entente, c'est : on prend les données, il y a des mouillages installés, on a pris les données et à partir de la fin de la mission, ils en ont une copie, tout le monde en a une copie, ils font ce qu'ils veulent avec, nous autres, on fait ce qu'on veut avec.* (Patrice, professeur en physique).

Dans ces derniers cas, le partage de données devient plus qu'une affaire strictement scientifique. Tant les chercheurs que leurs partenaires publics et privés doivent arrimer leurs objectifs et convenir d'ententes formelles.

#### 4.4.2 La spécialisation des tâches entre chercheurs

Le régime de sous-traitance peut prendre place à l'intérieur même d'une équipe de recherche, comme le montre le cas de Patrice, un des chercheurs que nous avons interviewés.

En tant que physicien, Patrice est appelé à collecter les données « noyau » (ou « core data »). Tel que mentionné précédemment, ces données physiques sont d'un grand intérêt pour plusieurs autres chercheurs (chimistes, biologistes, etc.). Compte tenu de l'importance critique de ses données pour le travail de ses collègues, Patrice et son équipe ont convenu d'une entente particulière avec *ArcticNet*. Cette entente prévoit de financer adéquatement la cueillette des variables physiques en échange du partage et de la distribution rapide de ces données à la communauté d'*ArcticNet*.

*[...] eux [les biologistes et les chimistes] ont besoin de la physique, [ils] ont besoin des propriétés physiques de l'eau de mer, ils ont besoin des courants, ils ont besoin de tout ça, donc je travaille un petit peu comme [...] océanographe de service, physicien de service pour fournir des données aux autres disciplines qui en ont absolument besoin.* (Patrice, professeur en physique).

La logique ainsi décrite par notre interviewé place donc l'équipe de physiciens dans une position de sous-traitant ayant pour « client » *ArcticNet*.

#### 4.4.3 L'encadrement formel des échanges

Alors que les deux précédents régimes s'appuient sur la confiance et la réciprocité, le régime de sous-traitance se démarque par son recours aux contrats, aux ententes, aux engagements officiels, etc. Dans les termes de Milanovic, Pontille et Cambon-Thomsen, il s'agirait d'un accord « contractuel » plutôt qu'un accord « moral » ou « conventionnel » (2007: 27).

Les échanges de données demeurent encadrés par un dispositif de documents écrits et signés. La nécessité de coordonner l'action d'acteurs hétérogènes (scientifiques, gouvernements, entreprises privées), d'explicitier les normes, les pratiques tacites de chacun et de traduire leurs langages respectifs rend le recours aux ententes presque inévitable. *A fortiori*, ces échanges impliquent souvent le transfert de ressources tangibles telles que des subventions et des infrastructures, lesquelles ne circulent pas sans traces écrites. À titre d'exemple, une interviewée nous explique comment le partage de données est étroitement encadré par un accord lié à l'obtention d'un permis.

*There is an agreement. For example, in one community that I'm working with, you have to have a research licence to do research in that community. As part of that agreement, you then have to return the digital files of the interviews as well as a transcriptions for them to have and then use, whenever they feel they need to. As well, participants obviously sign a consent form and then they indicate whether they want the transcript back to themselves personally as well.*  
(Émilie, étudiante en sciences humaines).

Dans le régime de sous-traitance, la contribution à des bases de données est généralement obligatoire. Elle est encadrée par des dispositions contractuelles et constitue une exigence à laquelle les chercheurs ne peuvent se soustraire, comme l'explique cette professionnelle de recherche :

*Une des exigences de ce projet-là était justement de construire une base de données commune. [...] C'était une prémisse de la subvention, c'était qu'ils devaient partager, ils devaient mettre publiques leurs données. [...] Là, on est rendu dans la fin du projet [...] c'est la fin des subventions. Donc présentement, les chercheurs,*

*ils doivent rentrer les données dans cette base de données-là pour qu'on puisse ensuite la rassembler et ça va être quelque chose qui va après ça être public. [...] Moi, c'est ça, j'essaie d'avoir tous les chercheurs qui nous soumettent les données, chose qui est pas évidente [léger rire].* (Tatiana, professionnelle de recherche en biologie).

En somme, le régime de sous-traitance regroupe les échanges de données qui s'effectuent dans le cadre de relations contractuelles et normées. Ces arrangements peuvent faire intervenir des acteurs externes au milieu scientifique et impliquer des échanges de ressources tangibles ou intangibles (données, permis, argent, subventions, infrastructures, etc.).

#### 4.5 Au-delà des régimes

Au cours des dernières pages, nous avons explicité les caractéristiques principales des régimes d'échange à la lumière des résultats de notre enquête auprès des chercheurs d'*ArcticNet*. Dans les prochaines lignes, nous présentons brièvement deux éléments du partage de données qui ont été abordés par les interviewés, mais qui ne correspondent à aucun des trois régimes d'échange en particulier.

##### 4.5.1 Le refus de partager

L'analyse révèle un des points aveugles de notre appareillage théorique. En effet, la théorie des régimes d'échange qualifie avec succès les échanges quand ils ont lieu, mais reste muette sur les situations où l'on refuse de partager les données. Pourtant, tant les études antérieures (Beaulieu, 2003; Campbell *et al.*, 2002) que nos entrevues montrent qu'il s'agit d'une réalité.

*[...] souvent les scientifiques, c'est la compétition, qui va publier le premier la meilleure chose. Donc, je pense que ça amène beaucoup de compétition puis il n'y a pas des échanges de données. Il y a beaucoup de portes fermées dans le domaine des échanges de données scientifiques. [...] En fait, des fois, j'ai rencontré des portes fermées dans mon propre labo [...].* (Edith, étudiante biologie).

*Ils [des collègues] ne veulent pas les donner à tout le monde [les données], tout de suite, ils veulent attendre... Ça, je peux comprendre. Il y en a d'autres qui, par principe, refusent de donner leurs données, mais y'en a de moins en moins* (Patrice, professeur en physique).

Le refus de partager s'avère donc être un enjeu concret dans la recherche sur l'Arctique et une question récurrente dans les travaux sur le partage de données. Cette question est traitée plus en profondeur au chapitre suivant.

#### 4.5.2 Les sanctions aux mauvais collaborateurs

La question des sanctions imposées aux « mauvais collaborateurs » (pour reprendre les mots de nos interviewés) est un autre aspect du partage de données qui retient notre attention, sans toutefois que nous puissions l'associer à un régime d'échange en particulier. Il s'agit des situations où l'un des partenaires manque à ses engagements ou agit contrairement aux normes de l'échange. Ces « mauvais collaborateurs » peuvent faire l'objet de mesures punitives, comme le discrédit :

*Il y a plutôt des ententes. Parce que si le gars le fait pas [se soumettre aux termes de l'entente], moi, je vais écrire à tous ceux qu'il connaît (Patrice, professeur en physique).*

Ces sanctions peuvent aussi prendre la forme de refus de partage subséquent :

*Si la personne à qui appartenait les données tombe sur un manuscrit publié, il y a rien à faire [...]. C'est sûr que bon, les gens qui ont fait ça une fois, après... leur accès aux données est beaucoup plus restreint ou surveillé (Paul, professeur en biologie).*

Comme dans le cas du refus de partage, les sanctions administrées aux « mauvais collaborateurs » semblent jouer un rôle important dans les échanges de données.

Maintenant que nous avons exposé nos résultats de recherche, nous les discutons plus en détail dans le prochain chapitre.



## CHAPITRE V

### DISCUSSION

Comme nous l'avons montré dans le dernier chapitre, les pratiques de partage de données sont bel et bien plurielles. Des échanges informels entre collègues jusqu'aux contrats signés entre chercheurs et partenaires privés, la circulation des données scientifiques revêt différentes formes. En mobilisant la typologie des régimes d'échange de Milanovic *et al.* (2007), nous sommes parvenus à mieux saisir les caractéristiques de ces pratiques et à explorer les contextes sociaux et scientifiques dans lesquels elles s'insèrent au sein d'*ArcticNet*. La présentation des résultats étant effectuée, nous sommes désormais en mesure de pousser l'analyse plus loin, de poser des constats plus généraux et d'émettre quelques hypothèses.

Le présent chapitre est divisé en trois sections, chacune d'entre elles représentant une des trois conclusions importantes de notre recherche. Premièrement, nous discutons de l'importance des relations humaines dans les échanges de données et des implications possibles sur la conception et le développement de bases de données. Deuxièmement, nous nous penchons plus en détail sur la relation qui s'établit lors de l'échange de données et sur les nouveaux rôles ou figures que revêtent les scientifiques dans ce contexte. Troisièmement, nous terminons le chapitre en discutant des limites de la théorie des régimes d'échange et en proposant un nouveau complément théorique.

#### 5.1 L'importance des relations de collaboration

Nous avons souligné la pluralité et la diversité des pratiques de partage de données. Différents acteurs sont impliqués dans différentes formes d'échange à des degrés variables de formalisation. Cette pluralité des pratiques étant bien évidente, nous en venons à nous poser la question inverse : qu'est-ce que tous ces échanges ont en commun?

### 5.1.1 La condition au partage : une relation de collaboration

Peu importe dans quel régime il se situe, nos résultats montrent que tout partage de données implique d'abord l'existence d'une relation de collaboration. Si un tel constat semble aller de soi, il n'en découle pas moins des implications très intéressantes. Avant même de faire circuler les données par courriel, par CD, par FTP ou par une base de données, les chercheurs communiquent entre eux intensément. Qu'il s'agisse de planifier une publication commune, de régler les détails d'un contrat ou d'apporter des précisions quant aux méthodes d'échantillonnages utilisées, les partenaires de l'échange de données établissent, renégocient ou renforcent une relation de collaboration.

Dans le régime de coopération, la dimension relationnelle est manifeste. Les interviewés nous parlent d'échange avec des collègues de longue date, qu'ils connaissent bien. Dans le régime de don, une relation préalable n'est pas requise pour partager des données, mais elle doit s'établir à travers la communication : « Il n'y a pas vraiment d'entente formelle, de choses signées, mais c'est des échanges de e-mails. » (Paul, professeur de biologie). Enfin, le régime de sous-traitance est beaucoup plus formalisé, mais il demeure que les partenaires de l'échange doivent établir une relation professionnelle significative. Dans tous les cas, l'importance de la confiance est palpable. Comme le font remarquer Milanovic, Pontille et Cambon-Thomsen, « la confiance est un élément essentiel de chacun des régimes d'échange, mais elle s'y exprime de manière différente [...] » (2007: 26).

En soulignant leurs ancrages relationnels, nous sommes amenés à recadrer notre perception des pratiques d'échange de données. En effet, il importe de considérer à la fois le transfert des données et les relations dans lesquelles il prend place. Encore ici, cette constatation semble triviale, mais pourtant elle nous permet d'émettre de nouvelles hypothèses sur les dispositifs techniques de partage de données.

### 5.1.2 Les technologies et les relations de collaboration

En tenant pour acquis la dimension éminemment relationnelle du partage des données scientifiques, nous sommes amenés à examiner comment les infrastructures de partage (comme les bases de données) et leurs discours d'accompagnement intègrent cette dimension. En d'autres termes, quelle place accorde-t-on aux relations entre chercheurs dans les projets de développement de bases de données?

Notre recension des écrits montre que la dimension relationnelle est presque totalement évacuée des discours d'accompagnement des bases de données en ligne (Anonyme, 2010; Arzberger *et al.*, 2004a, 2004b; Casey, 2003; Hey, Tansley et Tolle, 2009; National Research Council, 1997; Wellcome Trust, 2003; Wouters et Reddy, 2003). Le modèle de partage proposé correspond plutôt à l'idée de contributions individuelles à une grande base de données, laquelle peut ensuite être consultable à loisir. Autrement dit, l'action de faire don des données et l'action de les recevoir sont ici séparées. On omet tout bonnement de traiter des contacts interpersonnels qui s'établissent entre les scientifiques.

Pourtant, comme nous l'avons mentionné dans le chapitre précédent, le partage des données semble s'effectuer principalement de manière décentralisée, de pairs à pairs. Lors de ces échanges, ce ne sont pas que les données qui sont partagées, mais aussi des interactions riches, des négociations ou des précieuses indications sur la qualité des données.

*Celles-là [les données d'ancrages], c'est plus délicat. On les envoie quand même, mais en général, on dit aux gens « Ok, je te l'envoie, mais moi, j'ai un étudiant qui fait ça, ça, ça et ça » (Patrice, professeur en physique).*

En somme, les projets de base de données en ligne et les discours les accompagnant révèle une conception dépersonnalisée et l'instrumentalisée des échanges de données. Certes, une infrastructure centralisée comporte des avantages indéniables, grâce aux capacités d'archivage, d'organisation, de standardisation et de diffusion des données qu'elle offre. Cependant, en faisant fi de l'importance des relations interpersonnelles de collaboration entre les chercheurs, les promoteurs des bases de données prennent le pari risqué de soutenir une collaboration intense sans réelles assises relationnelles.

### 5.1.3 L'apport de l'informatique sociale

Les travaux en informatique sociale nous offrent une clé de lecture intéressante pour comprendre les discours d'accompagnement des bases de données. En effet, Kling (1999) remarque que les discours triomphalistes sont légion quand il s'agit de discuter des TIC. Il ne s'agirait pas d'une tendance neuve : «[...] les historiens ont noté que des déclarations utopiques et dystopiques ont été faites à propos de virtuellement tous les nouveaux dispositifs de communication [...]» (Boczkowski et Lievrouw, 2007a: 962). Loin de rester confinées à l'univers médiatique, ces visions utopistes des changements techniques influent directement sur

l'implantation de technologies dans les organisations. Les éditoriaux enthousiastes de *The Economist* sur le déluge de données (Anonyme, 2010) ou encore l'annonce d'un changement de paradigme dans les pages de *Nature* (Gilbert, 1991) ont certainement fait écho dans les universités et les corridors des agences de financement.

Pour Kling (1999), l'effet de cette rhétorique de la technologie révolutionnaire conduit souvent à l'adoption de dispositifs qui visent à remplacer l'action humaine, prétextant leur plus grande efficacité. Dans le même esprit, Breton (1987) soutient que le remplacement de l'action humaine par des machines est au cœur de l'histoire de l'informatique : « la conscience aiguë que l'homme est imparfait et que la solution à ce problème doit être recherchée dans la technique est sans doute l'un des héritages les plus importants que l'informatique ait reçus de son passé » (Breton, 1987: 143). En plus des questionnements éthiques que suscitent ces pratiques, Kling soutient que de tels projets se soldent par des échecs importants parce qu'ils n'ont pas su s'adapter convenablement aux contextes sociaux et culturels. Ainsi, le fait que les discours d'accompagnement des bases de données scientifiques omettent de traiter des échanges interpersonnels qui ont lieu quotidiennement est à la fois prévisible et inquiétant.

Pour les tenants de l'informatique sociale, l'approche à préconiser se doit d'éviter les déterminismes techniques et de conceptualiser les TIC comme des agents qui « restructurent les traitements de l'information déjà existants et les relations sociales qui les entourent » (Kling, Rosenbaum et Sawyer, 2005: 20). Dans les termes de la sociologie des usages, il s'agit de penser en terme d'évolution des usages anciens plutôt que de l'apparition spontanée d'usages inédits par l'introduction d'une nouvelle machine : « L'apparition de nouvelles pratiques se greffe sur le passé, sur des routines, sur des survivances culturelles [...] » (Mallein et Toussaint, 1994: 317). Du point de vue du design, cela peut se traduire par la nécessité de partir des pratiques existantes pour concevoir un outil, plutôt que de partir d'une technologie donnée et d'en imaginer les usages potentiels.

#### 5.1.4 Deux exemples d'intégration de la relation au dispositif

À partir du moment où l'on reconnaît l'importance de bâtir les dispositifs techniques à partir de l'observation des usages déjà existants, comment peut-on incorporer la dimension relationnelle de l'échange de données au sein de technologies telles que les bases de données? Les

deux exemples suivants illustrent bien deux manières distinctes de développer une infrastructure scientifique.

Un projet de base de données a provoqué une controverse importante dans les neurosciences au tournant des années 2000 (Anonyme, 2000a, 2000b; Beaulieu, 2003; Koslow, 2000; The Governing Council of the Organization for Human Brain Mapping, 2001). Le projet prévoyait exiger la contribution à la base de données pour publier dans les revues importantes de la discipline. L'opposition au projet a été très grande, notamment en raison des enjeux de propriété intellectuelle, d'éthique et de compétition entre chercheurs. Au-delà de ces considérations, le partage à travers une base de données centrale rendait la communication des facteurs contextuels et expérimentaux très difficile (Anonyme, 2000a). À l'opposé, le partage direct entre chercheurs facilitait cet échange d'information contextuelle. De même, en négociant les conditions de l'échange « un à un », il était davantage possible pour les chercheurs de concilier leurs besoins : protéger les données, assurer leur confidentialité et initier une relation de collaboration soutenue.

Le projet de base de données commune d'*ArcticNet* semble reconnaître la nature relationnelle du partage de données. En fait, la stratégie de développement du *Polar Data Catalogue* (PDC) a pour premier objectif de collecter les métadonnées de recherche. À travers l'interface web du PDC, on accède à des fiches qui détaillent la nature et le contexte des données : qui les a collectées, où, quand, avec quelles méthodes, dans le cadre de quel projet, etc. À la fin de chaque fiche, on retrouve l'adresse courriel du responsable du projet, avec l'invitation de le contacter si l'on est intéressé par ses données. Le reste du partage s'effectue ensuite directement entre les partenaires.

Cela dit, l'équipe de gestion du PDC désire se servir du dispositif en place pour archiver des données de recherche primaires. La phase 2 du projet du PDC prévoit archiver de plus en plus de données, tout en servant de guichet unique de soumission de métadonnées. En utilisant une approche graduelle (d'abord les métadonnées, ensuite les données elles-mêmes), le PDC semble davantage coller aux pratiques existantes et préserve la dimension relationnelle du partage de données.

Bien que quelque peu caricaturaux, ces deux exemples illustrent bien comment un même dispositif technique peut être intégré de manière très différente dans des communautés

scientifiques. Dans le premier cas, on envisage un renversement majeur des pratiques par l'utilisation d'un nouvel outil. Dans le deuxième cas, on prend le pari de partir des usages préexistants et d'intégrer la dimension relationnelle inhérente à l'échange de données dans le dispositif lui-même.

## 5.2 Une nouvelle figure : le chercheur-contributeur

S'il importe de considérer le contexte relationnel et les pratiques préexistantes de collaboration dans la conception des bases de données partagées, nous posons l'hypothèse de l'émergence d'une nouvelle figure du chercheur dans les milieux scientifiques : le chercheur-contributeur.

### 5.2.1 Comprendre le contexte et la qualité des données

Le fait que les chercheurs choisissent encore de procéder à des échanges un à un a de quoi surprendre. En effet, pourquoi s'encombrer de longues négociations, de dizaines de courriels et des incertitudes de la communication humaine quand il suffirait de consulter de grandes archives informatisées fournissant des données à loisir? Si l'information est disponible en quelques clics, pourquoi continue-t-on à faire appel à des intermédiaires?

Les chercheurs privilégient des partages « un à un » parce que les données sont généralement incomplètes. En effet, pour être en mesure d'user les données collectées par quelqu'un d'autre, d'en comprendre le contexte, d'en évaluer la qualité et les conditions de collecte, une interaction directe doit avoir lieu entre le producteur des données et la personne désireuse de les acquérir.

Un des indices les plus manifestes de ce besoin d'établir une relation est la préoccupation marquée des chercheurs interviewés quant à la qualité des données collectées par d'autres. Côté directement ses collègues reste une bonne manière de s'assurer de la qualité et de la représentativité des données partagées.

*C'est peut-être moins intéressant de partager des données avec quelqu'un qui a des chances d'introduire des erreurs parce qu'il [n']est pas consciencieux de ce qu'il fait. [...] Si t'as confiance en la personne, tu sais que cette personne, elle travaille bien, qu'elle est minutieuse, tout ça, t'as probablement plus confiance que les données qu'il va avoir obtenu vont être des bonnes représentations de la réalité. (Elena, étudiante en biologie).*

Certes, la présence de métadonnées complètes et précises peut grandement faciliter la compréhension et l'interopérabilité d'un jeu de données. Cependant, cette mesure s'avère toujours incomplète. Dans une étude ethnographique auprès de chercheurs en écologie, Zimmerman (2007, 2008) démontre que les informations fournies par les métadonnées ne sont pas suffisantes pour pouvoir comprendre le contexte de la saisie des données. Les sciences de l'environnement sont hautement contextuelles et les conditions particulières du terrain ont un impact majeur sur les données collectées. Ainsi, c'est à travers les expéditions sur le terrain, les expériences en laboratoire et les processus de socialisation (notamment durant les études supérieures) que les chercheurs développent une connaissance tacite de la qualité des données et de leurs contextes. Birnholtz et Bietz résument très bien cette situation :

*Ce degré de dépendance contextuelle signifie que de donner l'accès aux données requiert davantage que de simplement remettre un jeu de données. [...] Le récepteur doit souvent [...] passer du temps avec les chercheurs pour comprendre ce que les données signifient. [...] Ainsi, l'accès aux données et aux connaissances tacites qui les entourent devient une précieuse ressource [...]. (Birnholtz et Bietz, 2003: 343)*

### 5.2.2 Le chercheur-contributeur

Les chercheurs qui participent à un échange de données partagent bien plus que des données. Dans la conjoncture de surabondance des données que vivent de plus en plus de disciplines, l'habileté à ordonner de ce déluge informationnel est cruciale. C'est à travers leurs interactions que les chercheurs parviennent à s'échanger les connaissances cruciales pour bien comprendre et interpréter les jeux de données. Cette double circulation d'informations et de métainformations nous amène à proposer une nouvelle figure du chercheur, celle du chercheur-contributeur.

Le chercheur-contributeur est d'abord un chercheur qui s'inscrit dans une communauté, qui collecte des données, propose des hypothèses et des conclusions scientifiques. C'est lorsqu'il initie une relation de partage de données qu'il revêt un rôle de contributeur. En fournissant des données, il fournit aussi à ses collègues de précieuses indications sur la manière de les explorer, un peu à la manière dont le bibliothécaire guide les visiteurs dans une bibliothèque. Il détaille la topographie générale des données, les zones intéressantes et les territoires interdits (ex. : un étudiant de doctorat fait sa thèse sur cette partie des données à laquelle il ne faut pas toucher tant que la thèse n'est pas publiée). Il renseigne aussi ses collaborateurs sur le contexte particulier de

recueil des données, ainsi que sur les circonstances qui ont mené à leur production. En somme, il aide ses collègues à détecter le signal dans le bruit, à faire un peu d'ordre dans le chaos informationnel.

Nous pourrions dire que les chercheurs sont appelés à revêtir les fonctions de bibliothécaires ou de professionnels des sciences de l'information, sans pour autant avoir reçu la formation nécessaire.

### 5.2.3 Le chercheur-contributeur, figure du Web social?

Lorsqu'il revêt son rôle de contributeur, le chercheur agit de manière analogue à l'internaute idéalisé du Web 2.0, qui ne se contente pas de consommer un produit culturel, mais bien de contribuer à son évolution et à sa diffusion. En particulier, en pointant vers les données d'intérêt et en détaillant leur contexte, le chercheur effectue des actions très semblables aux usagers du Web social, qui soumettent des liens vers des contenus, les commentent et les distribuent dans leurs réseaux.

Pour Weinberger (2007), c'est à travers ce bourdonnement incessant d'échanges, de référencement, de commentaires, d'étiquetage et de discussion qu'émergerait l'ordre dans Internet. Contrairement aux classifications antérieures, telles que le système Dewey, l'ordonnement d'Internet serait perpétuellement en changement. Sous l'action des millions de contributions quotidiennes, la structuration d'Internet n'atteindrait jamais son point d'équilibre.

Si les actions des chercheurs-contributeurs ressemblent grandement à celles des internautes, les théorisations de Weinberger (2007), Proulx et Millerand (2010) sur le Web social peuvent-elles rendre compte des réalités scientifiques? Notre hypothèse serait d'affirmer que non. Malgré leurs ressemblances, les milieux scientifiques et les communautés du Web social s'inscrivent dans des contextes différents. Nous soutenons que la structure organisationnelle dans laquelle est enchâssée la recherche scientifique rend la contribution telle que nous l'observons dans le Web social beaucoup plus difficile. D'abord, la recherche fait intervenir un nombre relativement limité d'acteurs, ce qui contraste grandement avec la logique de la contribution dans le Web social qui s'appuie sur la force du nombre de ses usagers. Ensuite, l'exigence de la qualité et l'importance de la réputation dans les milieux savants rendent les contributions beaucoup plus difficiles. Enfin, malgré les efforts importants déployés par les chercheurs, les gouvernements et



les promoteurs de laboratoires, la science reste un lieu très décentralisé où la recherche et la collaboration s'effectuent principalement à l'échelle locale.

En l'absence d'infrastructures centrales capables d'agréger les contributions individuelles des chercheurs, le paradigme du Web social n'a sans doute pas encore franchi la porte des universités. Cependant, il existe de nombreux exemples d'initiatives scientifiques s'inspirant du modèle de la contribution du grand nombre. Wikipédia demeure probablement le cas le plus marquant, mais d'autres initiatives à plus petite échelle telles que Tela Botanica (Heaton, 2009) réussissent à faire cohabiter science et contribution populaire.

#### 5.2.4 Un contributeur invisible?

Pour terminer notre discussion sur la figure du chercheur-contributeur, nous associons celle-ci à la problématique du travail invisible dans la science (Shapin, 1989; Star et Strauss, 1999).

Dans son article *The Invisible Technician*, Shapin (1989) montre comment les travaux physiques, manuels et intellectuels accomplis par les techniciens sont historiquement occultés par les milieux scientifiques, même si ceux-ci sont absolument nécessaires à la production des connaissances. Pour Shapin, cette situation n'est pas uniquement le fait d'un passé lointain. Au contraire, dans les laboratoires d'aujourd'hui, les travaux et les apports des techniciens sont souvent relégués aux coulisses. L'article de Shapin a trouvé plusieurs échos, entre autres dans l'étude des dispositifs de collaboration (Star et Strauss, 1999).

En articulant notre théorisation émergente du chercheur-contributeur à celle du travail invisible de Shapin, nous posons l'hypothèse que le travail de contribution (commenter les données, prodiguer des indications, discuter des conditions d'utilisation) lors d'un échange de données est occulté. En d'autres termes, on évacue le rôle de contributeur du chercheur et l'on recadre la relation de collaboration en un transfert de données, en considérant les communications entre les chercheurs comme triviales et accessoires. Dans les discours d'accompagnement des bases de données scientifiques, l'omission des relations qui se tissent entre les chercheurs lors de collaboration représente un bon exemple de cette *mise en invisibilité*.

Dans la typologie du travail invisible proposée par Star et Strauss, il s'agirait d'une *mise en invisibilité* par la « disparition du travail dans le décor » (Star et Strauss, 1999: 15). Les efforts

importants investis pour ordonner, identifier, archiver, rendre accessible les données, maintenir la communication avec ses collaborateurs et effectuer la maintenance informatique sont considérés comme allant de soi et, au final, peu dignes de reconnaissance. Dans un contexte où les structures de récompenses du milieu scientifique sont grandement basées sur la « performance » en terme de publication, les efforts nécessaires à l'établissement de collaborations soutenues restent inapparents. La citation d'un interviewé rend très bien compte de la problématique :

*[...] c'est ça le gros défi, même si je reconnais tous les avantages, on reconnaît l'avantage de partager, l'avantage de rendre disponible, l'avantage de bien archiver, structurer, etc. J'en suis très conscient de tout ça, mais ça prend du temps, ça prend de l'énergie. Puis comme chercheur, la pression, c'est de publier à court terme, que tu sois très productif à court terme, donc on voit très bien les bénéfices à long terme, mais il n'y a pas un organisme subventionnaire qui me donne une subvention pour vingt ans. [...] Donc, pendant que je travaille pour structurer une base de données puis archiver puis rendre ça disponible aux autres, pendant ce temps-là, je n'écris pas d'articles scientifiques. (Pierre, professeur en écologie).*

En bref, les chercheurs sont impliqués dans un partage de données à la fois comme fournisseurs de données et comme contributeurs qui orientent leurs collaborateurs en leur prodiguant des conseils, des indications ou des restrictions sur l'utilisation de ces données.

### 5.3 Repenser les régimes d'échange

Au terme de cette discussion, il nous paraît nécessaire de faire un retour critique sur la théorie des régimes d'échange et de proposer de nouvelles pistes d'analyse et de développement. Nous soulignons d'abord deux limites théoriques au modèle proposé par Milanovic, Pontille et Cambon-Thomsen (2007) et nous terminons en présentant une contribution originale à la théorie des régimes d'échange pour dépasser ces limites.

#### 5.3.1 Une vision trop optimiste du partage de données

La première lacune importante de la théorie des régimes d'échange est le fait qu'elle ne traite que des échanges qui se déroulent bien. Nous l'avons déjà mentionné, le refus de partager ses données de recherche est monnaie courante dans le monde scientifique. Ce fait est à la fois confirmé dans nos entretiens et dans une étude quantitative à grande échelle (Campbell *et al.*, 2002). Pourtant, la théorie des régimes d'échange ne considère pas ces situations. En effet, elle ne

qualifie les relations d'échange que lorsqu'elles ont lieu, tout en restant muette sur les situations de refus de collaboration.

Le fait que la théorie des régimes d'échange ne porte son attention que sur les collaborations « réussies » a probablement un effet non négligeable dans la formation de nos outils méthodologiques et dans la conduite générale de notre recherche. En cherchant explicitement à rencontrer des chercheurs ayant déjà pris part à des échanges de données, nous sabordions nos propres chances d'examiner le phénomène du refus de partager dans toute sa complexité. N'ayant pas investigué davantage cette voie, les seuls témoignages collectés sur le refus de partager décrivent cette situation comme étant le fait de problématiques individuelles plutôt que systémiques :

*« Il y a des gens qui fondamentalement, ont des craintes ou ils font pas vraiment confiance aux autres donc leurs réflexes, c'est plutôt de garder l'information pour eux puis s'assurer que ça soit bien eux qui publient les résultats »* (Pierre, professeur en écologie).

Pourtant, tel que mentionné dans les dernières pages, il semble bien y avoir des facteurs contextuels importants qui rendent le partage difficile ou impossible.

En poussant cette réflexion un peu plus loin, nous sommes amenés à nous interroger sur les *a priori* normatifs de notre recherche. En effet, en mobilisant un cadre théorique qui rend compte uniquement des collaborations qui « fonctionnent » et qui néglige les situations où les échanges n'ont pas lieu, nous prenons parti pour le partage de données et taisons les voix discordantes. Cela nous conduit à dresser un portrait relativement partiel et partial du partage de données et nous prive d'éléments de réflexion. Il serait possible de remédier à ce point aveugle de la théorie des régimes d'échange en recadrant l'objet de notre analyse. Plutôt que d'étudier les régimes d'échange en tant que tels, nous serions appelés à « rediriger notre attention des produits vers les processus » (Jackson, Poole et Kuhn, 2002: 238). Ainsi, en détaillant comment et dans quelles conditions les régimes prennent place, nous ouvririons la porte aux cas où le partage ne s'établit pas. En d'autres mots, il s'agirait de passer d'un modèle statique des régimes à un modèle dynamique.

### 5.3.2 Pour un modèle dynamique des régimes d'échange

La théorie des régimes d'échange a été retenue pour notre cadre conceptuel à cause de sa remarquable clarté et de son grand potentiel pour l'analyse des pratiques d'échange. Bien entendu, la simplicité de cette catégorisation occulte nécessairement certains phénomènes et se révèle lacunaire dans certains cas. À notre avis, le principal manque à la théorie des régimes d'échange est de ne pas considérer explicitement les situations où les régimes changent, évoluent, se rompent, etc. En d'autres mots, la théorie des régimes d'échange pose les régimes comme étant relativement figés. Or, comme nous le mentionnions dans la dernière sous-section, le passage d'un modèle statique à un modèle dynamique nous permettrait de faire la lumière sur les processus de changement à l'œuvre dans les régimes, à la manière de Boltanski et Thévenot qui s'intéressent aux passages des acteurs d'un régime à un autre (Boltanski, 1990, 1993; Boltanski et Thévenot, 1991).

De même, nous sommes amenés à nous poser la question de la coexistence des régimes d'échange. À l'évidence, plusieurs régimes peuvent cohabiter dans un milieu donné. En fonction des projets, des collaborateurs, des données et des ressources, une même personne peut initier plusieurs régimes différents à la fois. Tous nos interviewés sont engagés dans au moins deux pratiques d'échange différentes. Comment conceptualiser les interactions entre les régimes d'échange?

### 5.3.3 Le passage d'un régime d'échange à un autre

Pour comprendre l'aspect dynamique des régimes d'échange, nous nous penchons sur les situations qui provoquent un changement d'un régime à un autre. Nos résultats relèvent au moins deux cas qui correspondent à cette définition.

Nous considérons d'abord les exemples de sanctions infligées aux « mauvais collaborateurs ». Les interviewés nous indiquent comment une mauvaise expérience de collaboration risque d'entacher la réputation d'une personne et fermer la porte à de possibles futurs échanges :

*Je vois quand même une limite à partager les yeux fermés. Ça va être un partage, mais peut-être que les gens qui se créent des réputations de bon collaborateur vont avoir plus facilement accès aux données des autres pis ceux qui se bâtissent une réputation de moins*

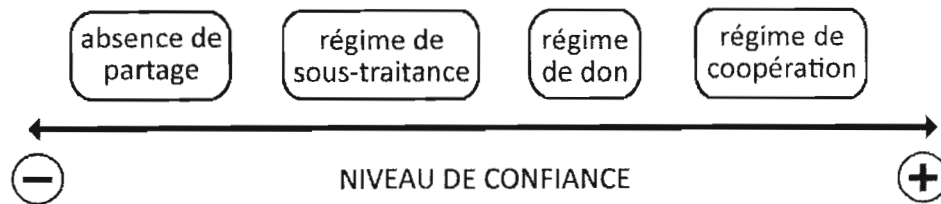
*bon collaborateur, ça sera plus difficile.* (Pierre, professeur en écologie).

La logique inverse est à l'œuvre quand des chercheurs se côtoient dans des collaborations répétées. La confiance qui s'y bâtit devient garante des échanges futurs :

*« [...] if you're requesting data and the group has been working together for a while and everyone has an understanding that you know that if... I don't worry that if someone gets data from they're not gonna credit me. »* (Elaine, étudiante en biologie).

Ces deux cas illustrent bien comment la confiance mutuelle entre les partenaires d'un échange peut être bâtie ou détruite. Dans les pages précédentes, nous avons souligné à quelques reprises le rôle prépondérant que joue la confiance dans tous les régimes d'échange. D'implicite et assurée dans le régime de coopération, elle devient de plus en plus codifiée dans le régime de sous-traitance. Enfin, quand elle est inexistante, aucun échange n'est possible.

À partir de ce constat, nous émettons l'hypothèse que le passage d'un régime à un autre s'effectue principalement en fonction du niveau de confiance réciproque des partenaires de l'échange. Ainsi, nous proposons la gradation suivante. Quand la confiance est absente, nous sommes dans des situations de refus de partage. Lorsque les partenaires gagnent un peu en confiance, ils peuvent échanger dans le régime de sous-traitance qui offre plusieurs garanties écrites et ou légales. Avec davantage d'assurance, les acteurs entrent dans le régime de don, dans lequel les obligations morales assurent une forme de réciprocité. Enfin, au degré de confiance maximale, les acteurs fonctionnent dans le régime de coopération et échangent librement des données.



**Figure 5.1** Gradation des régimes selon le niveau de confiance

Les changements dans le niveau de confiance pourraient donc initier des changements d'un régime à un autre, ou encore, dans les cas ultimes, d'un régime au refus de partager.

En revisitant notre problématique initiale, nous sommes amenés à considérer la situation sous un nouvel éclairage. En effet, il apparaît que la contribution aux bases de données en ligne peut être également conceptualisée sous la forme d'une « écologie » de régimes d'échange évoluant au gré de la confiance. L'histoire du succès de la base de données *Protein Data Bank* (PDB) montre bien comment la confiance de toutes les parties prenantes est lentement acquise au moyen de conventions, de traités et de mise en place de nouveaux dispositifs (Berman, Bourne et Westbrook, 2004). Après trente ans d'existence, la contribution dans la PDB semble être définitivement intégrée aux pratiques de la génomique et de la protéomique (Brown, 2003). Ainsi, les usagers et le dispositif de la PDB opèrent un déplacement graduel vers un état proche du régime de coopération, même si des règles générées dans les régimes antérieurs subsistent.

Le modèle proposé considère la confiance comme étant la dimension centrale du partage de données. L'importance de la confiance dans le modèle dynamique des régimes d'échange trouve écho dans certains travaux antérieurs sur le partage de données scientifiques. Entre autres, Beaulieu (2003) souligne que la confiance est nécessaire dans toutes les étapes du partage de données (cueillette, numérisation, transfert, archivage, retrait, etc.). Sans un niveau de confiance élevé, le partage est impossible : « Peu importe les bénéfices promis ou potentiels, la confiance doit être maintenue et entretenue pour que les pratiques de partage de données puissent se développer » (Beaulieu, 2003: 87). Beaulieu souligne aussi l'apport des infrastructures, des règles et des conventions scientifiques pour le développement de la confiance. Ces « mécanismes

générateurs de confiance » (*trust building mechanisms*) rendent le partage plus facile en assurant une part juste à tous les partenaires. C'est aussi à travers les règles et les infrastructures que les scientifiques peuvent asseoir leur légitimité auprès du public, en particulier dans les cas de données provenant de sujets humains.

En ce sens, ce modèle dynamique des régimes d'échange constitue une base de réflexion intéressante pour les promoteurs, concepteurs, administrateurs ou usagers d'infrastructures de collaboration scientifique. En conceptualisant les infrastructures et les politiques de partage comme des « mécanismes générateurs de confiance » plutôt que comme des outils triviaux, il est possible de développer une confiance véritable entre les scientifiques. De même, en prenant conscience du niveau de confiance des usagers et des régimes d'échange prévalents, les concepteurs de dispositifs peuvent prendre des décisions éclairées quant aux mesures à adopter.





## CONCLUSION

Pour clore ce mémoire, nous présentons une synthèse des apports principaux de notre recherche et traçons des pistes d'investigation futures. Pour ce faire, nous mettons d'abord en évidence les conclusions majeures de notre démarche. Ensuite, nous soulignons certaines limites de cette recherche. Enfin, nous terminons en proposant des pistes de recherche et d'analyse.

### 6.1 Le partage de données et la collaboration scientifique

#### 6.1.1 La pluralité des régimes d'échange

Notre recherche confirme l'existence d'une pluralité de pratiques en matière d'échange de données scientifiques. En effet, nous observons des pratiques aussi différenciées que la signature d'importants contrats de partenariat à des collaborations informelles et spontanées. Nous avons tenté de rendre compte de ces pratiques et de dégager les infrastructures, méthodes, contextes sociaux, communications réciproques, arrangements organisationnels et enjeux éthiques qui s'y rattachent.

Dans le régime de coopération, la proximité et la familiarité des acteurs rendent les échanges faciles et rapides (Milanovic, Pontille et Cambon-Thomsen, 2007). Sur notre terrain, ce régime prend forme entre des collègues qui participent à des expéditions communes, qui ont déjà collaboré par le passé ou qui travaillent ensemble au quotidien. Dans ce contexte, le recours à des bases de données intervient peu et représente plutôt une contrainte additionnelle pour les scientifiques.

Dans le régime de don, la relation d'échange est établie à travers un jeu réciproque de dons et de contre-dons (Milanovic, Pontille et Cambon-Thomsen, 2007). Sur notre terrain, ce régime se distingue du régime de coopération, car il peut faire intervenir des partenaires plus distants ou qui se connaissent peu. *Quid pro quo*, les données sont échangées contre d'autres ressources, le plus souvent l'ajout d'un nom d'auteur dans une publication. *A priori*, ce régime ne fait pas intervenir

directement l'usage d'infrastructures scientifiques (bases de données) et il prend généralement la forme d'un échange « un à un » entre deux scientifiques.

Dans le régime de sous-traitance, l'échange de données est assuré par une entente écrite dans laquelle un des partenaires rend accessibles ses données contre rétribution (Milanovic, Pontille et Cambon-Thomsen, 2007). Chez nos chercheurs interviewés, ce régime d'échange prend place dans les rares cas où ils collaborent avec des acteurs externes au milieu scientifique (entreprises, gouvernements). Les modalités de l'échange sont alors consignées dans des contrats ou ententes. Dans ce régime, la contribution de données dans une base est généralement exigée et stipulée dans les conditions du contrat.

Si cette pluralité de régimes d'échange semble triviale à première vue (il n'y a pas une mais plusieurs manières d'échanger et de partager ses données), la conception et la mise en œuvre d'infrastructures techniques aptes à rendre compte de cette diversité peuvent être des tâches très complexes. Pourtant, ces éléments nous apparaissent déterminants dans le succès des projets de bases de données scientifiques.

#### 6.1.2 Avant le partage, il y a une relation de collaboration

Le partage de données est avant tout une relation de collaboration. En effet, nous constatons la dimension primordiale que la relation de collaboration prend dans le partage. La confiance, la proximité, la réciprocité et les communications mutuelles sont cruciales dans le partage des données. Encore ici, le design et la gestion de projets d'outils collaboratifs devraient intégrer cette dimension relationnelle, autant pour des raisons d'efficacité (une plus grande probabilité de succès) que pour mieux s'intégrer aux pratiques déjà en place dans la communauté scientifique.

#### 6.1.3 Le chercheur-contributeur, une nouvelle figure à explorer

Nous posons l'hypothèse de l'émergence de la figure du chercheur-contributeur dans la collaboration scientifique. Nous distinguons une première forme de partage, le partage des données primaires, d'une deuxième forme de partage, le partage *d'informations* sur ces données. Les chercheurs s'échangent en effet bien plus que de simples données, ils s'échangent aussi des indications précises sur les conditions d'utilisation de ces données (à propos de leur pertinence, leur contexte de cueillette, leurs zones d'intérêts, leurs territoires interdits, etc.). Cet échange

relativement informel de « métadonnées » nous apparaît être un des éléments cruciaux du partage de données, et cela, surtout dans le présent contexte de surabondance informationnelle.

Certains travaux proposent des analyses analogues à celle que nous développons autour de la notion de chercheur-contributeur. Ainsi, Lenoir (1998) examine comment le travail des biologistes s'est transformé avec le développement des banques de données moléculaires. Par exemple, les biologistes ont développé des collaborations avec des mathématiciens et des informaticiens, tout en s'initiant eux-mêmes aux outils de gestion des données (voir aussi Fujimura, 1999; Fujimura et Fortun, 1996). Les bases de données étant, dans ce domaine, déjà nombreuses, des collaborations s'établissent autour de l'usage des logiciels d'analyse, des méthodes, des métadonnées, etc.

Dans le même sens, Gray *et al.* (2009) soutiennent que le travail des scientifiques est sur le point de connaître un changement épistémique majeur. L'extrême abondance de données de recherche conduit les chercheurs à penser le travail scientifique de plus en plus comme de la gestion d'informations et de connaissances. C'est dans le développement de modèles, d'algorithmes et d'outils que se situeraient désormais le cœur du travail et de la collaboration scientifique.

Nous sommes d'avis que la figure du chercheur-contributeur peut permettre d'envisager la conception et l'usage des infrastructures de collaboration différemment, en privilégiant un point de vue plus holistique de la collaboration et du partage de données scientifiques. En effet, si l'on veut établir une base de données en ligne qui regroupe un grand nombre de collaborateurs, il faut établir des mécanismes par lesquels les chercheurs peuvent s'échanger des informations, des conseils et des commentaires sur les données partagées. Comme Bowker (2000) le démontre, les données scientifiques sont loin d'être des unités discrètes et neutres, elles sont imprégnées d'une série de décisions théoriques, sociales, organisationnelles et politiques depuis leur cueillette jusqu'à leur mise en base. Dans ce contexte, l'échange de métainformations entre chercheurs est essentiel.

## 6.2 Les limites de la recherche

Nous consacrons les prochains paragraphes à la mise en évidence des limites et des lacunes de notre recherche.

### 6.2.1 Le biais positif de la théorie des régimes d'échange

Nous l'avons déjà mentionné, la théorie des régimes d'échange (Milanovic, Pontille et Cambon-Thomsen, 2007) ne considère que les situations d'échange lorsqu'elles ont lieu et lorsqu'elles se déroulent bien.

Il s'agit là d'une limite théorique importante. Premièrement, cela nous amène à porter une attention démesurée sur les cas de réussite tout en voilant les cas de non-partage, de collaboration conflictuelle ou de négociations difficiles. En concevant la collaboration et le partage de données comme systématiquement positives et sans accrocs, le portrait final finit par être déformé. En effet, les résultats de notre recherche et des études antérieures nous indiquent que le refus de partager ses données est monnaie courante dans les milieux scientifiques (Beaulieu, 2002, 2003; Campbell *et al.*, 2002).

Deuxièmement, l'insistance de la théorie des régimes d'échange sur les collaborations réussies nous empêche d'aborder en profondeur la question du refus de partager. Les implications méthodologiques sont majeures : dans le cadre de notre recherche, nous avons explicitement cherché à rencontrer des chercheurs qui partagent leurs données. Lors des entrevues, la question du refus de partager n'a été abordée qu'indirectement, lorsque les interviewés nous ont parlé de collègues plus réfractaires.

Par contre, l'incapacité de la théorie initiale à rendre compte des situations problématiques de partage ne nous empêche pas de recueillir des témoignages intéressants sur ces cas et d'en discuter dans ce mémoire. En effet, certains interviewés nous ont fait part de partages conflictuels et de leur contexte. Dans la discussion, nous proposons un modèle dynamique des régimes d'échange, qui inclut les situations de refus de partager et prévoit des changements de régime dans les cas conflictuels.

### 6.2.2 Les limites de nos appuis empiriques

Le design de notre recherche devant se conformer aux exigences d'un travail de recherche de maîtrise, nos incursions sur le terrain de recherche ont été limitées dans le temps. Notre principal corpus d'analyse, les entrevues semi-dirigées, représente environ cinq heures trente minutes d'enregistrement audio, ce qui est peu pour répondre aux questions importantes de la collaboration scientifique et du partage de données. Nos conclusions doivent donc être

considérées comme des hypothèses et des conceptualisations préliminaires plutôt que comme des théories confirmées.

Le peu d'appuis empiriques est un problème récurrent dans les études sur les bases de données et le partage en science. La majorité des travaux relevés dans la recension des écrits sont des études de cas isolées, guidées par des objectifs de recherche et des méthodologies disparates. Dans une publication antérieure, nous soulignons cette disparité des recherches et relevions le besoin d'une conceptualisation solide pour organiser durablement ce champ d'études (Gratton-Gagné et Millerand, à paraître). La présente recherche a pour objectif de contribuer à cette consolidation en reprenant la théorisation des régimes d'échange (Milanovic, Pontille et Cambon-Thomsen, 2007).

### 6.3 Les pistes de recherches futures

Pour terminer cet ouvrage, nous recommandons quatre avenues d'investigation que nous jugeons particulièrement porteuses.

#### 6.3.1 Les nouvelles formes de travail scientifique

Dans les dernières pages, nous avons émis l'hypothèse du « chercheur-contributeur » qui, d'une part partage ses données et d'autre part, oriente les futurs usagers des données dans le « déluge » informationnel. Dans ce contexte, les chercheurs sont de plus en plus amenés à intégrer des façons de faire et des pratiques propres aux bibliothécaires ou, plus largement, aux professionnels de l'information et de la documentation. Le travail des chercheurs s'hybriderait donc avec celui des professionnels de l'information. Si une telle hypothèse est avérée, nous soutenons qu'il importe de reconsidérer de manière importante la structure et l'organisation du travail scientifique : comment doit-on former les futurs chercheurs? Comment doit-on organiser les systèmes de récompense? Quels rôles les professionnels de recherche et les étudiants seront-ils amenés à jouer?

Il y a place à des études novatrices sur ces nouvelles formes de travail scientifique. Plusieurs perspectives peuvent permettre de creuser la question, notamment celle du travail invisible (Shapin, 1989; Star et Strauss, 1999), les approches plus critiques ou encore les études féministes. Plus de vingt ans après la parution de l'article *The invisible technician* de Steven Shapin (1989), le rôle, les tâches et les accomplissements des techniciens restent très peu

considérés et sous-documentés dans les études en science, technologie et société (STS). En joignant la problématique du partage de données à celle des rapports de travail dans les milieux scientifiques, il serait possible de développer des programmes de recherche fort prometteurs.

Ainsi, l'examen du travail concret que représentent le développement et l'entretien d'une base de données et l'observation de la façon dont ce travail est négocié, réparti et représenté au sein des équipes de recherche pourrait constituer un terrain extrêmement riche. À cet effet, des travaux récents mettent en lumière le « travail d'articulation » nécessaire à l'établissement d'infrastructures scientifiques (Baker et Millerand, 2007). En grande partie assumé par des gestionnaires d'information, ce travail d'articulation est bien souvent invisible et non prévu dans la planification et le développement des systèmes de collaboration (Baker et Millerand, 2007; Karasti et Baker, 2004; Star et Strauss, 1999).

De même, il y a lieu de s'interroger sur les conséquences épistémologiques et éthiques de l'évolution du travail scientifique vers le modèle de la gestion et de l'analyse d'immenses bases de données (Borgman, Wallis et Enyedy, 2007; Gray *et al.*, 2009; Hey, Tansley et Tolle, 2009; Lenoir, 1998). Beaulieu (2001) note par exemple que le passage des laboratoires expérimentaux (*wet labs*) aux laboratoires informatiques (*dry labs*) en neurosciences provoque des changements importants dans la mesure où l'obtention, la manipulation et le traitement de jeux de données informatisées constituent désormais les tâches principales du travail scientifique. Les neurologues craignent ainsi de voir leur science se transformer en une série d'exercices managériaux plutôt que d'être guidée par des questions de recherche motivantes (Beaulieu, 2001). Dans cette perspective, l'étude critique du rôle des technologies dans la transformation des milieux scientifiques représente une avenue de recherche incontournable.

### 6.3.2 L'étude quantitative du partage de données

Si la majorité des travaux sur le partage de données mobilisent des méthodologies qualitatives, nous soutenons qu'il y aurait place à des travaux de nature quantitative, qui, au moyen de sondages ou d'analyses scientométriques, permettraient de dégager les tendances lourdes de l'évolution du partage de données scientifiques. Jusqu'à présent, Campbell *et al.* (2002) et Brown (2003) sont les seuls chercheurs à avoir traité la question du partage de données par des méthodes quantitatives. Or, le déploiement à grande échelle de cyberinfrastructures ou l'adoption de politiques scientifiques pour favoriser le partage requièrent d'autres études

quantitatives. Il existe plusieurs occasions de recherche dans ce secteur. Entre autres, la citation de jeux de données dans les publications pourrait fournir des indices sur l'usage des bases de données selon les disciplines.

### 6.3.3 Développer la théorie des régimes d'échange

La théorie des régimes d'échange de Milanovic, Pontille et Cambon-Thomsen (2007) n'a pas eu beaucoup d'écho dans la sphère scientifique. Un des objectifs secondaires de notre étude est de redonner un second souffle à cet effort de théorisation, qui comme nous l'avons vu, se révèle tout à fait approprié pour comprendre les échanges qui ont lieu dans le réseau *ArcticNet*. À notre avis, les recherches futures gagnent à explorer de nouveau cette théorie.

À la fois simple, flexible et heuristique, nous croyons que la théorie des régimes d'échange est en mesure de qualifier bien plus que des modes de partage de données. Les relations de collaboration et de partage qui se tissent dans les espaces du Web social (Millerand, Proulx et Rueff, 2010) ou encore dans des réseaux professionnels gagneraient à être étudiées sous l'angle des régimes d'échange. À titre d'exemple, l'une des stratégies d'affaire et de développement actuelles est la « coopétition », soit la coopération *et* la compétition simultanées entre plusieurs entreprises (Brandenburger et Nalebuff, 1998). Dans ce contexte où les membres de deux organisations sont appelés à travailler avec leurs concurrents, le cadre conceptuel des régimes d'échange peut être employé pour révéler les formes de collaboration et de partage.

### 6.3.4 Agir et réfléchir sur les dispositifs de collaboration

La dernière piste d'investigation que nous souhaiterions voir se développer est l'établissement de programmes de recherche-action ou recherche-intervention dans le domaine des outils collaboratifs en science. En prenant appui sur le corpus grandissant d'études empiriques et théoriques sur ces technologies de collaboration, des chercheurs-praticiens peuvent mettre en place des outils et des processus qui parviennent à développer durablement une culture de collaboration et de partage dans les milieux scientifiques.

Notre expérience de terrain l'a montré, le partage de données et le recours aux technologies de collaboration sont des enjeux importants pour les chercheurs de l'Arctique. Or, notre recension des écrits indique que les chercheurs d'autres disciplines sont confrontés aux mêmes défis. Les sciences de la santé et les sciences humaines gagneraient à être étudiées au moyen d'une

recherche-action, notamment en raison des critères éthiques et méthodologiques qui balisent leur travail. Enfin, nous soutenons que les études en communication et les études en STS sont particulièrement bien outillées pour ce type de recherche. Leurs apports multidisciplinaires, leur problématisation de la technique et leur bagage théorique peuvent être d'une grande aide pour l'analyse et le développement de pratiques collaboratives en science (Boczkowski et Lievrouw, 2007b, 2007a).

Les dernières pages ont montré comment ce mémoire ne constitue pas un point final, mais bien une amorce pour des travaux futurs sur la question des infrastructures scientifiques, du partage de données et des relations de travail dans la sphère scientifique. À notre avis, ces enjeux sont cruciaux, notamment dans nos sociétés où l'information et les savoirs sont placés au centre des arènes politiques, économiques et culturelles. Nous l'avons déjà évoqué, la recherche sur l'Arctique est au cœur d'enjeux géopolitiques internationaux. Dans ce contexte, une réflexion véritable sur le rôle sociétal des sciences et des technologies s'impose. Le présent travail se veut donc une modeste contribution en ce sens, laquelle aura, espérons-le, de nombreuses suites.

- - -

- - - -

- - -



## BIBLIOGRAPHIE

- Anderson, J. A., et G. Baym. 2004. «Philosophies and Philosophic Issues in Communication, 1995-2004». *Journal of Communication*. vol. 54, no 4, p. 589-615.
- Anonyme. 2000a. «A debate over fMRI data sharing». *Nature Neuroscience*. vol. 3, no 9, p. 845-846.
- . 2000b. «Whose scans are they, anyway?». *Nature*. vol. 406, no 6795, p. 443-443.
- . 2002. «regime». In *Dictionary of the Social Sciences*, Craig Calhoun. New York: Oxford University Press. En ligne.  
<<http://www.oxfordreference.com/views/ENTRY.html?subview=Main&entry=t104.e1420>>.
- . 2008. «Police probe theft of MoD laptop». BBC News. En ligne.  
<<http://news.bbc.co.uk/2/hi/7197045.stm>>. Consulté le 25 mars 2010.
- . 2010. *Data, data everywhere. A special report on managing information*. Londres: The Economist.
- ArcticNet. 2010a. *Rapport annuel 2009-2010*. En ligne.  
<<http://www.arcticnet.ulaval.ca/pdf/media/AR2010.pdf>>.
- . 2010b. «Research Phase II (2008-2011)». En ligne.  
<<http://www.arcticnet.ulaval.ca/research/phase2.php>>. Consulté le 20 septembre 2010.
- Arzberger, P., P. Schroeder, A. Beaulieu, G. C. Bowker, K. Casey, L. Laaksonen, D. Moorman, P. Uhler et P. Wouters. 2004a. «An International Framework to Promote Access to Data». *Science*. vol. 303, no 5665, p. 1777-1778.
- . 2004b. «Promoting Access to Public Research Data for Scientific, Economic, and Social Development». *Data Science*. vol. 3, p. 135-152.

- Atkins, D. E. 2003. «Revolutionizing Science and Engineering Through Cyberinfrastructure». Report of the NSF blue-ribbon Advisory Panel on Cyberinfrastructure: National Science Fondation
- Auray, N. 2004. «La régulation de la connaissance : arbitrage sur la taille et gestion aux frontières dans la communauté Debian». *Revue d'économie politique*, no 113, p. 160-182.
- . 2006. «Les configurations de marché du logiciel et le renouvellement du capitalisme». In *L'économie des conventions, méthodes et résultats*, François Eymard-Duvernay. Paris: La Découverte.
- Auray, N., M. Hurault-Plantet, M. Poudat et B. Jacquemin. 2009. «La négociation des points de vue: une cartographie sociale des conflits et des querelles dans le Wikipédia francophone». *Réseaux*, no 154, p. 15-50.
- Baker, K. S., et F. Millerand. 2007. «Articulation Work Supporting Information Infrastructure Design: Coordination, Categorization, and Assessment in Practice». *40th Hawaii International Conference on System Sciences (HICSS 07)* (Hawaii, 2-6 janvier).
- Baym, N. K. 1995. «The Emergence of Community in Computer-Mediated Communication». In *CyberSociety. Computer-Mediated Communication and Community*, Steven G. Jones, p-138-163. Thousand Oaks: Sage.
- Beaulieu, A. 2001. «Voxels in the Brain: Neuroscience, Informatics and Changing Notions of Objectivity». *Social Studies of Science*. vol. 31, no 5, p. 635-680.
- . 2002. «Withholding of data among academic geneticists». *Journal of the American Medical Association*. vol. 287, no 15, p. 1939-1940.
- . 2003. «Research woes and new data flows». In *Promise and practice in data sharing*, Paul Wouters et Peter Schroder, p. 65-88. Amsterdam: NIWI-KNAW.
- Berg, M., et E. Goorman. 1999. «The contextual nature of medical information». *International Journal of Medical Informatics*. vol. 56, p. 51-60. Nlm.
- Berman, H. M., P. E. Bourne et J. Westbrook. 2004. «The Protein Data Bank: A Case Study in Management of Community Data». *Current Proteomics*, no 1, p. 49-57.
- Bijker, W. E., T. P. Hughes et T. Pinch. 1994. *The Social Construction of Technological Systems*. Coll. «New Directions in the Sociology and History of Technology». Cambridge: MIT Press.

- Birnholz, J. P., et M. J. Bietz. 2003. «Data at work: supporting sharing in science and engineering». In *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*: ACM.
- Boczkowski, P. J., et L. A. Lievrouw. 2007a. «Bridging Communication Studies and Science and Technology Studies: Scholarship on Media and Information Technologies». In *The Handbook of Science and Technology Studies*, 3e, Edward J. Hackett, Olga Amsterdamska, Michael Lynch et Judy Wajcman, p. 949-978. Cambridge, MA: MIT Press.
- . 2007b. «Bridging Communication Studies and Science and Technology Studies: Scholarship on Media and Information Technologies». In *Proceedings of the International Communication Association Conference* (24-28 mai).
- Bohlin, I. 2004. «Communication Regimes in Competition: The Current Transition in Scholarly Communication Seen through the Lens of the Sociology of Technology». *Social Studies of Science*. vol. 34, no 3, p. 365-391.
- Boltanski, L. 1990. *L'amour et la justice comme compétences : trois essais de sociologie de l'action*. Coll. «Leçons de choses». Paris: Métailié.
- . 1993. «Dissémination ou abandon: la dispute entre amour et justice. L'hypothèse d'une pluralité de régimes d'action». In *La Théorie de l'action : le sujet pratique en débat*, Louis Quéré, p. 235-259. Paris: CNRS.
- Boltanski, L., et L. Thévenot. 1991. *De la justification : les économies de la grandeur*. Coll. «NRF essais». Paris: Gallimard.
- Borgman, C. L., J. C. Wallis et N. Enyedy. 2007. «Little science confronts the data deluge: habitat ecology, embedded sensor networks, and digital libraries». *International Journal of Digital Library*. vol. 7, no 1, p. 17-30.
- Bos, N., A. S. Zimmerman, J. Olson, J. Yew, J. Yerkie et al. 2007. «From shared databases to communities of practice: A taxonomy of collaboratories». *Journal of Computer-Mediated Communication*. vol. 12, no 2.
- Bowker, G. C. 2000. «Biodiversity datadiversity». *Social Studies of Science*. vol. 30, no 5, p. 643-683.
- . 2006. *Memory Practices in the Sciences*. Cambridge, MA: MIT Press.

- Brandenburger, A., et B. Nalebuff. 1998. *Co-opetition*, 1st pbk. New York ; Toronto: Currency Doubleday.
- Breton, P. 1987. *Histoire de l'informatique*. Coll. «L'histoire des sciences». Paris: Editions La Découverte.
- Breton, P., et S. Proulx. 2002. *L'Explosion de la communication à l'aube du XXIe siècle*. Montréal: Boréal.
- Brown, C. 2003. «The changing face of scientific discourse: Analysis of genomic and proteomic database usage and acceptance». *Journal of the American Society for Information Science and Technology*. vol. 54, no 10, p. 926-938.
- Burk, D. 2007. «Intellectual property in the context of e-science». *Journal of Computer-Mediated Communication*. vol. 12, no 2.
- Burrell, G., et G. Morgan. 1979. *Sociological Paradigm and Organization Analysis*. Londres: Heinemann.
- Callon, M. 1994. «Society in the Making: The Study of Technology as a Tool for Sociological Analysis». In *The Social Construction of Technological Systems*, 5, Wiebe E. Bijker, Thomas P. Hughes et Trevor Pinch. Cambridge, MA: MIT Press.
- Campbell, E. G., B. R. Clarridge, M. Gokhale, L. Birenbaum, S. Hilgartner, N. A. Holtzman et D. Blumenthal. 2002. «Data Withholding in Academic Genetics: Evidence From a National Survey». *Journal of the American Medical Association*. vol. 287, no 4, p. 473-480. En ligne. <<http://jama.ama-assn.org/cgi/content/abstract/287/4/473>>.
- Carlson, S., et B. Anderson. 2007. «What are data? The many kinds of data and their implications for data re-use». *Journal of Computer-Mediated Communication*. vol. 12, no 2.
- Casey, K. 2003. «Issues of electronic data access in biodiversity». In *Promise and practice in data sharing*, Paul Wouters et Peter Schroder, p. 41-64. Amsterdam: NIWI-KNAW.
- Cassier, M. 1998. «Le partage des connaissances dans les réseaux scientifiques : l'invention de règles de "bonne conduite" par les chercheurs». *Revue française de sociologie*. vol. 39, no 4, p. 701-720. Consulté le 18 avril 2009.
- Chivers, T. 2007. «Data loss crisis spreads to the NHS». *The Telegraph* (Londres), 23 décembre. En ligne. <<http://www.telegraph.co.uk/news/uknews/1573465/Data-loss-crisis-spreads-to-the-NHS.html>>.

- Coopmans, C. 2006. «Making mammograms mobile: Suggestions for a sociology of data mobility». *Information, Communication & Society*. vol. 9, no 1, p. 1-19. Consulté le December 30, 2008.
- Corcuff, P. 1996. «Ordre institutionnel, fluidité situationnelle et compassion. Les interactions au guichet de deux CAF». *Recherches et prévisions*, no 45.
- . 1998. «Justification, stratégie et compassion: Apport de la sociologie des régimes d'action». *Correspondances (Bulletin d'information scientifique de l'Institut de Recherche sur le Maghreb Contemporain)*, no 51.
- Corti, L. 2000. «Progress and Problems of Preserving and Providing Access to Qualitative Data for Social Research - The International Picture of an Emerging Culture». *Qualitative Social Research*. vol. 1, no 3.
- de Certeau, M. 1990. *L'Invention du quotidien*, Nouv. ed. Coll. «Collection Folio/essais». Paris: Gallimard.
- Deslauriers, J.-P., et M. Kérisit. 1997. «Le devis de recherche qualitative». In *La recherche qualitative. Enjeux épistémologiques et méthodologiques*, Jean Poupard, Jean-Pierre Deslauriers, Lionel-Henri Groulx, Anne Laperrière, Robert Mayer et Alvaro P. Pires, p. 85-111. Montréal: Gaëtan Morin Éditeur.
- Dictionary of the Social Sciences. 2002. «regime». In *Dictionary of the Social Sciences*, Craig Calhoun. New York: Oxford University Press. En ligne.  
<<http://www.oxfordreference.com/views/ENTRY.html?subview=Main&entry=t104.c1420>>.
- Dietrich, M. R., et B. H. Tambasco. 2007. «Beyond the boss and the boys: Women and the division of labor in *Drosophila* genetics in the United States, 1934-1970». *Journal of the History of Biology*, no 40, p. 509-528.
- Dortier, J.-F. 2004. «Foucault, Michel». In *Le dictionnaire des sciences humaines*, p. 257-259. Auxerre: Sciences Humaines.
- Drake, M. A. 2003. «Free Public Access to Science - Will It Happen?». In «*Information Today, Inc.*». En ligne.  
<<http://newsbreaks.infotoday.com/nbreader.asp?ArticleID=16666>>. Dernière modification le 7 juillet 2003. Consulté le 16 décembre 2008.

- Eckersley, P., G. F. Egan, S. Amari, F. Beltrame, R. Bennett, J. G. Bjaalie, T. Dalkara, E. De Schutter, C. Gonzalez, S. Grillner, A. Herz, K. P. Hoffmann, I. P. Jaaskelainen, S. H. Koslow, S. Y. Lee, L. Matthiessen, P. L. Miller, F. M. da Silva, M. Novak, V. Ravindranath, R. Ritz, U. Ruotsalainen, S. Subramaniam, A. W. Toga, S. Usui, J. van Pelt, P. Verschure, D. Willshaw, A. Wrobel et Y. Tang. 2003. «Neuroscience data and tool sharing: a legal and policy framework for neuroinformatics». *Neuroinformatics*. vol. 1, no 2, p. 149-165.
- Flichy, P. 1991. *Une histoire de la communication moderne*. Coll. «Histoire des sciences». Paris: La Découverte.
- Fujimura, J. H. 1999. «The Practices of Producing Meaning in Bioinformatics». *Sociology of Science Yearbook*. vol. 21, no 1, p. 49-87.
- Fujimura, J. H., et M. Fortun. 1996. «Constructing Knowledge across Social Worlds: The Case of DNA Sequence Databases in Molecular Biology». In *Naked Science: Anthropological Inquiry into Boundaries, Power and Knowledge*, Laura Nader, p. 160–173. New York: Routledge.
- Gilbert, W. 1991. «Towards a paradigm shift in biology». *Nature*. vol. 349, no 6305, p. 99.
- Glaser, B. G., et A. L. Strauss. 1967. *The discovery of grounded theory : strategies for qualitative research*. New York: Aldine.
- Goldenberg, A. 2010. «La participation dans les communautés épistémiques: don ou contribution». In *Web social : mutation de la communication*, Florence Millerand, Serge Proulx et Julien Rueff, p. 217-230. Québec: Presses de l'Université du Québec.
- Granjon, F. 2004. «De quelques éléments programmatiques pour une sociologie critique des usages sociaux des TIC». In *les rapports société-technique du point de vue des sciences de l'homme et de la société* (mai 2004).
- Granovetter, M. S. 1973. «The Strength of Weak Ties». *The American Journal of Sociology*. vol. 78, no 6, p. 1360-1380.
- Gratton-Gagné, O. 2009. «Les défis de la collaboration : un bref panorama des études sur les bases de données dans les communautés scientifiques.». In *8e colloque interuniversitaire en communication* (13 mars - inédit).
- Gratton-Gagné, O., et F. Millerand. à paraître. *Des infrastructures de collaboration pour le partage des données scientifiques : bibliographie commentée et résumée*. Montréal: Centre interuniversitaire de recherche sur la science et la technologie (CIRST).

- Gray, J., A. J. G. Hey, S. Tansley et K. M. Tolle. 2009. «Jim Gray on eScience: A Transformed Scientific Method». In *The fourth paradigm : data-intensive scientific discovery*, Anthony J. G. Hey, Stewart Tansley et Kristin Michele Tolle, p. xvii-xxxi. Redmond, Wash.: Microsoft Research.
- Guba, E. 1990. *The paradigm dialogue*. Newbury Park: Sage Publications.
- Heaton, L. 2009. «Les savoirs botaniques en réseau: le travail scientifique des professionnels et des amateurs à Tela Botanica». *Colloque «Figures de l'usager en contexte d'innovation»* (Montréal, QC).
- Hey, A. J. G., S. Tansley et K. M. Tolle. 2009. *The fourth paradigm : data-intensive scientific discovery*. Redmond, Wash.: Microsoft Research.
- Hilgartner, S. 1995. «Biomolecular Databases: New Communication Regimes for Biology?». *Science Communication*. vol. 17, no 2, p. 240-263.
- Hine, C. 2006. «Databases as scientific instruments and their role in the ordering of scientific work». *Social Studies of Science*. vol. 36, no 2, p. 269-298.
- Hughes, T. P. 1994. «The Evolution of Large Technological Systems». In *The Social Construction of Technological Systems*, 5, Wiebe E. Bijker, Thomas P. Hughes et Trevor Pinch. Cambridge, MA: MIT Press.
- Jackson, M. H., M. S. Poole et T. Kuhn. 2002. «The Social Construction of Technology in Studies of the Workplace». In *Handbook of New Media*, Leah A. Lievrouw et Sonia Livingstone, p. 236-253. Londres: Sage Publications.
- Jones, E., B. Watson, J. Gardner et C. Gallois. 2004. «Organizational Communication: Challenges for the New Century». *Journal of Communication*. vol. 54, no 4, p. 722.
- Jouët, J. 2000. «Retour critique sur la sociologie des usages». *Réseaux*, no 100, p. 487-521.
- Karasti, H., et K. S. Baker. 2004. *The Long-Term Information Management Trajectory: Working to Support Data, Science and Technology*. San Diego: Scripps Institution of Oceanography Technical Report, University of California San Diego.
- Katzen, M. 1986. «The application of computers in the humanities: A view from Britain». *Information Processing & Management*. vol. 22, no 3, p. 259-267.

- Keil, B. 1990. «Cooperation between databases and scientific community». In *Methods in Enzymology*, p. 50-60: Academic Press.
- Kling, R. 1999. «What is Social Informatics and Why Does it Matter?». *D-Lib Magazine*. vol. 5, no 1.
- . 2004. «The Internet and unrefereed scholarly publishing». *Annual Review of Information Science and Technology*. vol. 38, p. 591-631.
- Kling, R., et E. Callahan. 2003. «Electronic journals, the Internet, and scholarly communication». *Annual Review of Information Science and Technology*. vol. 37, p. 127-177.
- Kling, R., et H. Crawford. 1999. «From retrieval to communication: The development, use, and consequences of digital documentary systems». *Journal of the American Society for Information Science*. vol. 50, no 12, p. 1121-1122.
- Kling, R., et G. McKim. 1999. «Scholarly communication and the continuum of electronic publishing». *Journal of the American Society for Information Science*. vol. 50, no 10, p. 890-906.
- Kling, R., G. McKim et A. King. 2003. «A bit more to it: Scholarly communication forums as socio-technical interaction networks». *Journal of the American Society for Information Science and Technology*. vol. 54, no 1, p. 47-67.
- Kling, R., H. Rosenbaum et S. Sawyer. 2005. *Understanding and communicating social informatics : a framework for studying and teaching the human contexts of information and communication technologies*. Medford, N.J.: Information Today.
- Koslow, S. H. 2000. «Should the neuroscience community make a paradigm shift to sharing primary data?». *Nature Neuroscience*. vol. 3, no 9, p. 863-865.
- Laperrière, A. 1997. «La théorisation ancrée (grounded theory): démarche analytique et comparaison avec d'autres approches apparentées». In *La recherche qualitative. Enjeux épistémologiques et méthodologiques*, Jean Poupard, Jean Pierre Deslauriers, Lionel Henri Groulx, Anne Laperrière, Robert Mayer et Alvaro P. Pires, p. 309-340. Montréal: Gaëtan Morin Éditeur.
- Latour, B. 1987. *Science in action : how to follow scientists and engineers through society*. Cambridge, MA: Harvard University Press.



- . 1990. «Drawing things together». In *Representation in Scientific Practice*, M. Lynch et S. Woolgar, p. 19–68. Cambridge, MA: MIT Press.
- . 1999. «Circulating reference: Sampling the soil in the Amazon forest». In *Pandora's hope: Essays on the reality of science studies*, p. 24-79. Cambridge, MA: Harvard University Press.
- Latour, B., et S. Woolgar. 1979. *Laboratory life : the social construction of scientific facts*. Coll. «Sage». Beverly Hills: Sage Publications.
- Leeds-Hurwitz, W. 1992. «Forum Introduction: Social Approaches to Interpersonal Communication». *Communication Theory*. vol. 2, no 2, p. 131-139.
- Lelong, B., F. Thomas et C. Ziemlicki. 2004. «Des technologies inégalitaires? L'intégration de l'Internet dans l'univers domestique et les pratiques relationnelles». *Réseaux*. vol. 22, no 127-128, p. 141-180.
- Lenoir, T. 1998. «Shaping Biomedicine as an Information Science». In *Proceedings of the 1998 Conference on the History and Heritage of Science Information Systems*, Mary Ellen Bowden, Trudi Bellardo Hahn et Robert V. Williams, p. 27-45. Medford, NJ: Information Today. Consulté le 16 décembre 2008.
- Luo, A., et J. S. Olson. 2008. «How Collaboratories Affect Scientists from Developing Countries». In *Scientific Collaboration on the Internet*, Gary M. Olson, Ann Zimmerman et Nathan Bos. Cambridge, Mass.: MIT Press.
- Mallein, P., et Y. Toussaint. 1994. «L'intégration sociale des technologies d'information et de communication: une sociologie des usages». *Technologies de l'information et société*. vol. 6, no 4, p. 315-336.
- Mallein, P., Y. Toussaint et M. Bydlowski. 1984. *Teletel 3V, les adolescents et leur famille*: CEPS Centre d'études des pratiques sociales, Université des sciences sociales Grenoble-II.
- Mauss, M. 1950. «Essai sur le don - Forme et raison de l'échange dans les sociétés archaïques». In *Sociologie et anthropologie*. Paris: PUF.
- Mayer, R., et F. Ouellet. 1991. *Méthodologie de recherche pour les intervenants sociaux*. Boucherville: Gaëtan Morin Éditeur.
- Merton, R. K. 1974. *The sociology of science : Theoretical and empirical investigations*. Chicago: University of Chicago Press.

- Michener, W. K., J. W. Brunt, J. J. Helly, T. B. Kirchner et S. G. Stafford. 1997. «Nongeospatial Metadata for the Ecological Sciences». *Ecological Applications*. vol. 7, no 1, p. 330-342.
- Milanovic, F., D. Pontille et A. Cambon-Thomsen. 2007. «Biobanking and Data Sharing: a Plurality of Exchange Regimes ». *Genomics, Society and Policy*. vol. 3, no 1, p. 17-30.
- Miles, M. B., et A. M. Huberman. 1984. *Qualitative data analysis : a sourcebook of new methods*. Beverly Hills: Sage.
- Miller, K. 2005. «Philosophical Foundations: What Is Theory?». In *Communication Theories, Perspectives, Processes and Contexts*, 2e, p. 20-31: McGraw Hill.
- Millerand, F. 2007. «Des infrastructures d'information pour les scientifiques : analyse des processus d'innovation sociotechnique au sein d'un projet d'infrastructure de recherche.». Description du projet de recherche: FQRSC
- Millerand, F., S. Proulx et J. Rueff. 2010. *Web social : mutation de la communication*. Québec: Presses de l'Université du Québec.
- Nachi, M. 2003. *Ethique de la promesse. L'agir responsable*. Coll. «Philosophies». Paris: PUF.
- National Research Council. 1997. *Bits of power: Issues in global access to scientific data (Free Executive Summary)*. Washington, DC: National Academy Press. En ligne. <[www.nap.edu/catalog.php?record\\_id=5504#toc](http://www.nap.edu/catalog.php?record_id=5504#toc)>.
- National Science Board. 2005. *Long-Lived Digital Data Collections Enabling Research and Education in the 21st Century*. Arlington, VA: National Science Foundation.
- O'Reilly, T. 2005. «What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software». En ligne. <<http://oreilly.com/web2/archive/what-is-web-20.html>>. Consulté le 18 mars 2010.
- Ogien, A., et L. Quéré. 2005a. *Le vocabulaire de la sociologie de l'action*. Coll. «Vocabulaire de...». Paris: Ellipses.
- . 2005b. «Régimes d'action». In *Le vocabulaire de la sociologie de l'action*, p. 105-110. Paris: Ellipses.

- Olson, G. M., et J. S. Olson. 2000. «Distance Matters». *Human-Computer Interaction*, vol. 15, no 2, p. 139-178.
- Olson, G. M., A. Zimmerman et N. Bos. 2008. *Scientific Collaboration on the Internet*. Coll. «Acting with Technology». Cambridge, Mass.: MIT Press.
- Paillé, P. 1994. «L'analyse par théorisation ancrée». *Cahiers de recherche sociologique*, no 23.
- Phillips, D. C. 1992. *The Social Scientist's Bestiary: A Guide to Fabled Threats to and Defences of Naturalistic Social Science*. Oxford: Pergamon.
- Pontille, D., F. Milanovic, E. Rial-Sebbag et A. Cambon-Thomsen. 2006. *Le vivant à l'épreuve des collections : biobanques et bases de données en recherche biomédicale*. Toulouse: Centre d'Étude et de Recherche Travail, Organisation, Pouvoir et Épidémiologie et analyses en Santé Publique.
- Poupart, J. 1997. «L'entretien de type qualitatif: considération épistémologiques, théoriques et méthodologiques». In *La recherche qualitative. Enjeux épistémologiques et méthodologiques*, Jean Poupart, Jean-Pierre Deslauriers, Lionel-Henri Groulx, Anne Laperrière, Robert Mayer et Alvaro P. Pires, p. 173-210. Montréal: Gaëtan Morin Éditeur.
- Proulx, S. 2010. «La figure du contributeur : la création de valeur dans le contexte du capitalisme informationnel». *Figures de l'usager en contexte d'innovation* (Montréal, 29 janvier 2010). CIRST.
- Proulx, S., et F. Millerand. 2010. «Le Web social, au carrefour de multiples questionnements». In *Web social : mutation de la communication*, Florence Millerand, Serge Proulx et Julien Rueff, p. 13-30. Québec: Presses de l'Université du Québec.
- Quéré, L. 1993. *La Théorie de l'action : le sujet pratique en débat*. Paris: CNRS.
- Radhakrishna, R. 2010. «Minding the Gap of Global Health Inequity: Dr. Hans Rosling on Visual Statistics, Trend Analysis, and Storytelling ». *Global Pulse*, vol. 6, no 1, p. 1-8. En ligne. <<http://www.globalpulsejournal.com/spfi10a.html>>.
- Sawyer, S., K. Crowston, R. T. Wigand et M. Allbritton. 2003. «The Social Embeddedness of Transactions: Evidence from the Residential Real-Estate Industry». *Information Society*, vol. 19, no 2.

- Schroder, P. 2003. «Digital Research Data as Floating Capital of the Global Science System». In *Promise and practice in data sharing*, Paul Wouters et Peter Schroder, p. 7-11. Amsterdam: NIWI-KNAW.
- Shapin, S. 1989. «The Invisible Technician». *American Scientist*. vol. 77, no 6, p. 554-563.
- Star, S. L. 1988. «Introduction: The Sociology of Science and Technology». *Social Problems*. vol. 35, no 3, p. 197-205.
- Star, S. L., et K. Ruhleder. 1994. «Steps towards an ecology of infrastructure: complex problems in design and access for large-scale collaborative systems». In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*: ACM.
- . 1996. «Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces». *Information Systems Research*. vol. 7, no 1, p. 111-134.
- Star, S. L., et A. Strauss. 1999. «Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work». *Computer Supported Cooperative Work*. vol. 8, no 1-2, p. 9-30.
- Summers, D. 2007. «Personal details of every child in UK lost by Revenue & Customs». *The Guardian* (Londres), 20 novembre 2007. En ligne.  
<<http://www.guardian.co.uk/politics/2007/nov/20/economy.personalfinancenews>>.  
Consulté le 14 avril 2009.
- Talja, S., et H. Maula. 2003. «Reasons for the use and non-use of electronic journals and databases». *Journal of Documentation*. vol. 59, no 6, p. 673-691.
- The Governing Council of the Organization for Human Brain Mapping. 2001. «Neuroimaging Databases». *Science*. vol. 292, no 5522, p. 1673-1676.
- Weinberger, D. 2007. *Everything is miscellaneous : the power of the new digital disorder*. New York: Times Books.
- Wellcome Trust. 2003. *Sharing Data from Large-Scale Biological Research Projects: A System of Tripartite Responsibility. Report of a meeting organized by the Wellcome Trust and held on 14–15 January 2003 at Fort Lauderdale, USA*. En ligne.  
<[www.genome.gov/Pages/Research/WellcomeReport0303.pdf](http://www.genome.gov/Pages/Research/WellcomeReport0303.pdf)>.
- Wouters, P., et C. Reddy. 2003. «Big science data policies». In *Promise and practice in data sharing*, Paul Wouters et Peter Schroder, p. 13-40. Amsterdam: NIWI-KNAW.

- Wouters, P., K. Vann, A. Scharnhorst, M. Ratto, I. Hellsten, J. Fry et A. Beaulieu. 2007. «Messy Shapes of Knowledge - STS Explores Informatization, New Media, and Academic Work». In *The Handbook of Science and Technology Studies*, 3e, Edward J. Hackett, Olga Amsterdamska, Michael Lynch et Judy Wajcman, p. 319-352. Cambridge, MA: MIT Press.
- Zimmerman, A. S. 2007. «Not by metadata alone: the use of diverse forms of knowledge to locate data for reuse». *International Journal on Digital Library*. vol. 7, no 1, p. 5-16.
- . 2008. «New knowledge from old data - The role of standards in the sharing and reuse of ecological data». *Science Technology & Human Values*. vol. 33, no 5, p. 631-652.